Journal of Medical Internet Research

Journal Impact Factor (JIF) (2022): 7.4
Volume 12 (2010), Issue 1 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Short Paper

Original Paper

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

Thomas L Webb[1]; Judith Joseph[2]; Lucy Yardley[2]; Susan Michie[3]

[1]Department of Psychology, University of Sheffield, Western Bank, Sheffield, UK

[2]School of Psychology, Shackleton Building, University of Southampton, Highfield, Southampton, UK

[3]Research Department of Clinical, Educational & Health Psychology, University College London, 1-19 Torrington Place, London, UK

Corresponding Author:
Thomas L Webb
Department of Psychology
University of Sheffield
Western Bank, Sheffield, S10 2TN
UK
Phone: +44 114 222 6516
Fax: +44 114 276 6515
Email: t.webb@sheffield.ac.uk

Abstract

Background: The Internet is increasingly used as a medium for the delivery of interventions designed to promote health behavior change. However, reviews of these interventions to date have not systematically identified intervention characteristics and linked these to effectiveness.

Objectives: The present review sought to capitalize on recently published coding frames for assessing use of theory and behavior change techniques to investigate which characteristics of Internet-based interventions best promote health behavior change. In addition, we wanted to develop a novel coding scheme for assessing mode of delivery in Internet-based interventions and also to link different modes to effect sizes.

Methods: We conducted a computerized search of the databases indexed by ISI Web of Knowledge (including BIOSIS Previews and Medline) between 2000 and 2008. Studies were included if (1) the primary components of the intervention were delivered via the Internet, (2) participants were randomly assigned to conditions, and (3) a measure of behavior related to health was taken after the intervention.

Results: We found 85 studies that satisfied the inclusion criteria, providing a total sample size of 43,236 participants. On average, interventions had a statistically small but significant effect on health-related behavior ($d_+ = 0.16$, 95% CI 0.09 to 0.23). More extensive use of theory was associated with increases in effect size ($P = .049$), and, in particular, interventions based on the theory of planned behavior tended to have substantial effects on behavior ($d_+ = 0.36$, 95% CI 0.15 to 0.56). Interventions that incorporated more behavior change techniques also tended to have larger effects compared to interventions that incorporated fewer techniques ($P < .001$). Finally, the effectiveness of Internet-based interventions was enhanced by the use of additional methods of communicating with participants, especially the use of short message service (SMS), or text, messages.

Conclusions: The review provides a framework for the development of a science of Internet-based interventions, and our findings provide a rationale for investing in more intensive theory-based interventions that incorporate multiple behavior change techniques and modes of delivery.

(J Med Internet Res 2010;12(1):e4) doi:10.2196/jmir.1376

KEYWORDS

Internet; intervention; behavior change; meta-analysis; review

Introduction

...without a scientific underpinning, the field [of Internet interventions] may flounder [1] [LM Ritterband and DF Tate]

In June 2009 an estimated 25% of the world's population had access to the Internet, with estimates in Europe and North America being considerably higher (50% and 74%, respectively) [2]. Researchers in the field of health promotion have been quick to capitalize on the exponential growth of the Internet, and over the past decade, an increasing number of interventions designed to promote changes in health behavior have been delivered via the Internet [1,3]. For example, "Happy Ending" is a 54-week Internet-based intervention designed to promote smoking abstinence [4,5]. This intervention involves over 400 contact emails that direct participants to a different webpage each day, supplemented by interactive voice response (IVR) and short message service (SMS) monitoring and prompts. Other Internet-based interventions, however, simply involve embedding a short planning exercise within an online lifestyle survey [6,7].

Quantitative reviews of Internet-based interventions report positive–albeit highly variable and often small–effects on behaviors such as physical activity, tobacco use, exercise, and so on [8-12]. However, previous reviews have not systematically coded the characteristics of each Internet-based intervention and computed the effect size associated with each [1]. The limited analyses of this kind that have been conducted suggest that this approach may provide insight into the characteristics of effective versus less effective interventions. For example, Portnoy et al [8] coded whether the intervention included information, motivation, or behavioral skill components. The findings suggested that the inclusion of motivational components (eg, cost-benefit analyses) actually weakened the impact of the interventions. Since the publication of the review by Portnoy et al [8], a comprehensive taxonomy of behavior change techniques has been published [13], along with a method for assessing the extent to which behavioral interventions are theory-based [14]; both these developments permit more sophisticated coding of intervention content. Thus, the primary aim of the present review was to use these new coding schemes to identify the characteristics of effective Internet-based interventions. A secondary aim was to develop a coding scheme for the different modes by which Internet-based interventions are delivered (eg, via scheduled access to an advisor or automated feedback) and to link different modes of delivery to effect size.

How Can the Characteristics of Internet-based Interventions Be Conceptualized?

Three intervention characteristics may influence the impact on behavior [15-18]: (1) the theoretical basis of the intervention, (2) the behavior change techniques used, and (3) the mode of delivery.

Theoretical Basis and Use of Theory and Predictors

Theoretical basis refers to the theory or theories used to develop the intervention. For example, in an effort to promote physical activity, Spittaels et al [19] directed participants to a website that presented a tailored message based on the theory of planned behavior [20]. In contrast, Carr et al [21] used social cognitive theory [22] to develop a physical activity intervention that could be delivered via the Internet. Theory can inform interventions in a number of different ways, from identifying theoretical constructs to be targeted (eg, attitude, self-efficacy) or mechanisms underlying particular behavior change techniques (eg, vicarious learning in modeling), to selecting participants most likely to benefit (eg, people with particularly negative attitudes). Despite assertions that use of theory leads to more effective interventions [23-27], there is debate over the importance of theory [28,29], and at present it is unclear whether and how use of theory influences intervention effectiveness, particularly in relation to Internet-based interventions [1]. A large review of HIV-prevention interventions reported that use of theory was positively related to extent of behavior change [30], but this finding was simply based on whether or not theory was cited. Although this is an important step in the right direction, it would be useful to know how different uses of theory impact on the effectiveness of interventions and whether more extensive use of theory leads to larger effects than less extensive use. Michie and Prestwich [14] have developed a reliable coding scheme to assess the different ways that behavioral interventions employ theory; use of this coding scheme permits the present review to investigate these important questions.

Behavior Change Techniques

Behavior change techniques refer to the specific strategies used in the intervention to promote behavior change. For example, some interventions designed to promote smoking abstinence prompt barrier identification and problem solving (eg, [31]), whereas other interventions prompt participants to monitor their behavior (eg, [32]). In order to identify techniques contributing to effectiveness across interventions and to ensure that effective interventions can be replicated, it is crucial that standardized definitions of the techniques included in behavior change interventions are used and linked to intervention effectiveness [33]. With this in mind, the present review used the taxonomy of behavior change techniques developed by Abraham and Michie [13] to code the content of the interventions.

Mode of Delivery

The interventions in the present review were delivered via the Internet. The effects of this primary mode of delivery can be estimated by examining studies that compare similar materials presented via the Internet versus other modes, such as print [34,35]. Internet-based interventions can, however, differ substantially in their specific mode of delivery. For example, content can be delivered in a more or less interactive manner [36,37]. Interventions may also employ supplementary delivery modes (eg, SMS messaging, email, telephone, or videoconferencing) that may influence effectiveness. To our knowledge, no coding scheme exists for assessing the mode with which Internet-based interventions are delivered. Existing coding schemes developed for systematic reviews of non-Internet interventions [38] are not suitable because they focus on the physical manner in which participants received the intervention (eg, one-to-one or group) and the nature of the

person delivering the intervention (eg, health educator or trained facilitator). Therefore, the present review developed a new coding scheme for assessing mode of delivery in Internet-based interventions and used it to understand how each mode influences the effectiveness of the intervention.

The Present Review

The present review sought to investigate which characteristics of Internet-based interventions were associated with effectiveness. By so doing, we answer the important applied and theoretical questions: Which theories should researchers draw on in developing interventions? How can theory best be used to inform Internet-based interventions? What behavior change techniques are effective when employed via the Internet? Is the mode by which the intervention is delivered important?

Method

Selection of Studies

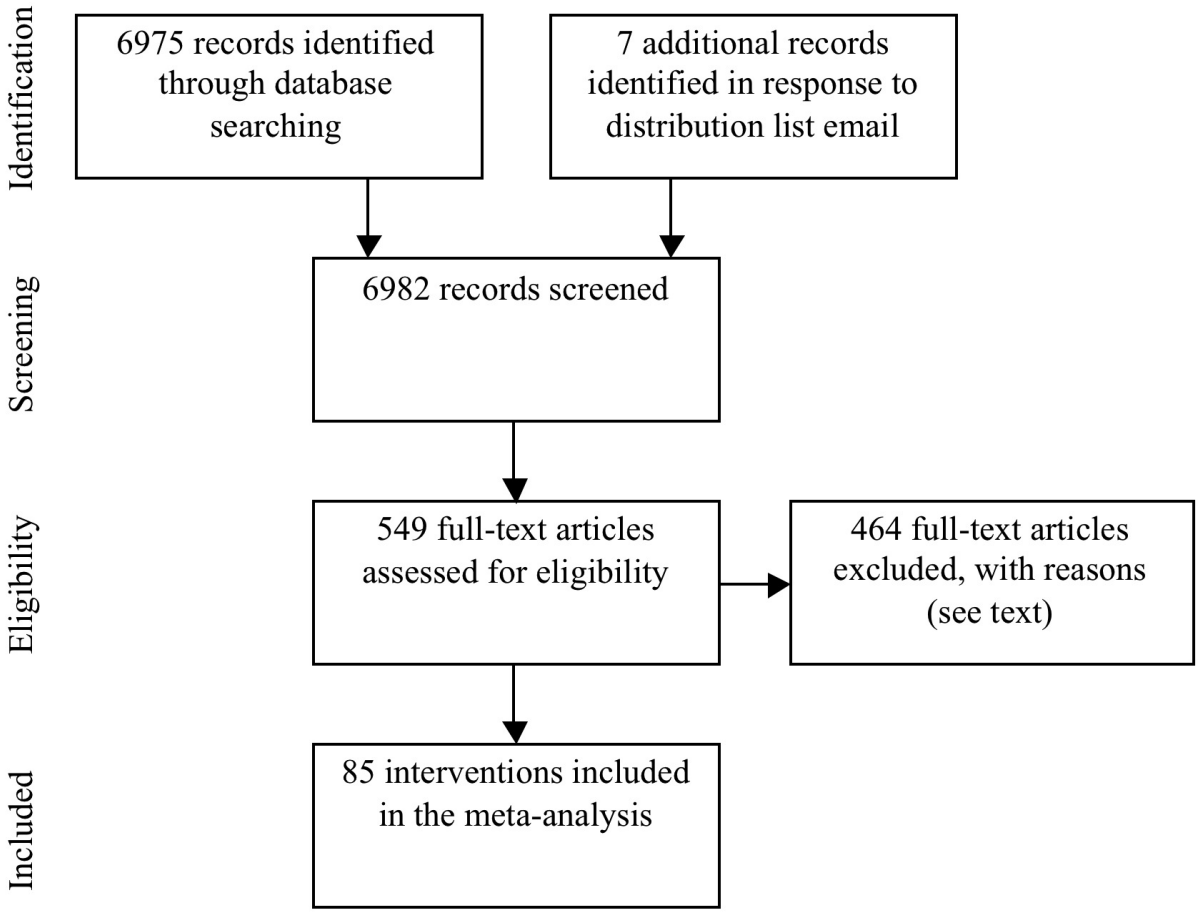
Identification and Screening

In July 2008 we conducted a computerized search using ISI Web of Knowledge, which covers a number of databases including Web of Science conference proceedings (1900-), BIOSIS Previews (1985-), and Medline (1950-). We used the following search terms: Web-based, Internet, digital, online, technolog*, computer, treatment, RCT, trial, intervention, behavio* change. (The asterisk automatically truncates the term such that, for example, technolog* will also find technology, technologies, etc). Studies had to include one or more of the search terms in the title. We also sent an email to the distribution list of the European Health Psychology Society to request unpublished research. There were three inclusion criteria for the meta-analysis. First, the primary components of the intervention must have been delivered via the Internet (not including CD-ROMs, SMS messaging, or other computer applications). Second, the studies must have involved random assignment of participants to a treatment group that received an Internet-based intervention and a comparison group that received either a control intervention or no intervention. Finally, a measure of behavior related to health must have been taken after the intervention. We did not include studies that only measured symptoms (eg, anxiety, depression), cognitions (eg, attitudes, intentions), outcomes presumed to be the consequence of behavioral changes (eg, weight loss, blood glucose levels), or behaviors unrelated to health (eg, use of literature services).

Eligibility and Inclusion

Figure 1 shows the flow of information through the different phases of the review. We assessed 549 full-text articles for eligibility. Of these, 140 studies (26%) were rejected because the study did not include a measure of behavior related to health (eg, [39]), 97 studies (18%) were rejected because the primary components of the intervention were not delivered via the Internet (eg, [40]), 88 studies (16%) were rejected because they did not report intervention effects (typically, these were reviews or protocol descriptions, eg, [41]), 84 studies (15%) were rejected because they did not include a control group (eg, [42]), 20 studies (4%) were rejected because computers were used only to tailor information that was presented in a non-computerized format (eg, [43]), 17 studies (3%) were rejected because they reported additional effects of an intervention already included in the review (eg, [44]), 8 studies (1%) were rejected because intervention effects were reported in a manner that did not permit computation of an effect size (eg, [45]). For these studies, it was decided not to estimate effect sizes based on the significance levels reported because all the effects for which full information was not available were reported as non-significant. Assuming zero difference (d = 0.00) for these effects could systematically underestimate effect sizes associated with particular intervention characteristics. Finally, 5 studies (1%) were rejected because participants were not randomly allocated to conditions (eg, [46]), and 4 studies (1%) were rejected because the manuscripts were not written in English (eg, [47]). In total, 85 reports of Internet-based interventions met the inclusion criteria for the review. Multimedia Appendix 1 presents the characteristics and effect sizes associated with each intervention.

Figure 1. Flow of information through the different phases of the review (adapted from [48])



Calculation of Effect Sizes for the Effect of Internet-Based Interventions on Health-Related Behavior

The effect size for post-intervention behavior differences between the conditions was calculated in line with Cochrane recommendations [49]. Specifically, the longest follow-up was selected wherever possible. For example, where Brendryen et al [5] followed up smokers at 3, 6, and 12 months, the 12-month data was included in the review. Where studies examined more than one behavior (such as Williamson et al's study [50] of weight loss behaviors in which exercise, overeating, and avoidance of fattening foods were measured), the effect sizes within the study associated with different behaviors were meta-analysed in their own right prior to inclusion in the main dataset. This procedure captures the richness of the data and does not prioritize one outcome over another (eg, effects on dietary outcomes and effects on physical activity are considered equally important), while also maintaining the independence of samples that is central to the validity of meta-analysis [51]. Intention-to-treat analyses were used wherever possible. Following Portnoy et al [8], where studies employed more than one comparison condition, we selected the most passive comparison condition for ease of interpretation. For a detailed discussion of considerations relating to choice of comparison conditions, see Danahar and Seeley [53].

Coding of Intervention Characteristics

Use of Theory and Predictors

The coding scheme developed by Michie and Prestwich [14] was used to code how theory and predictors (constructs that are not explicitly linked to a theory by the authors but are targeted for intervention because they predict behavior) were used in the design of the interventions. Items 1 through 6 of the coding scheme identify whether theory or predictors are mentioned and whether they are used to select recipients for the intervention, to select or develop intervention techniques, or to tailor intervention techniques to participants. Items 7 through 11 examine whether intervention techniques are explicitly linked to theory-relevant constructs or predictors and, conversely, whether theory-relevant constructs or predictors are linked to intervention techniques. Items 12 through 17 were not evaluated in the present review because they do not pertain to use of theory in developing the intervention. These items focus on methodological issues (randomization and measurement quality) and whether theory was refined on the basis of outcomes. Where the theoretical basis of the experimental intervention was identical to that of the comparison intervention (eg, [34]), the intervention was coded as not having a theoretical basis that could explain differences in effect size between the conditions.

In addition to considering each use of theory separately, we also summed items 1 through 11 to create an overall "use of theory"

score that could be used to evaluate whether more extensive use of theory leads to larger effects than less extensive use. In a slight change to the published recommendations, item 8 ("At least one, but not all, intervention techniques are explicitly linked to at least one theory-relevant construct/predictor") was coded as "yes" if item 7 ("All intervention techniques are explicitly linked to at least one theory-relevant construct/predictor") was coded as "yes." Similarly, item 11 ("At least one, but not all, theory-relevant constructs/predictors are explicitly linked to at least one intervention technique") was coded as "yes" if item 10 ("All theory-relevant constructs/predictors are explicitly linked to at least one intervention technique") was coded as "yes." This ensured that when we created the "use of theory" score, reports that linked, for example, all theoretical constructs with intervention techniques, were also credited as linking some theoretical constructs with intervention techniques.

Theoretical Basis

Interventions were coded as having a particular theoretical basis only if the theory was used to develop the intervention techniques (item 5 of the coding scheme of Michie and Prestwich [14]) rather than theory being simply mentioned (item 1).

Behavior Change Techniques

The behavior change techniques used in the interventions were coded using an augmented 40-item version [52] of the 26-item taxonomy developed by Abraham and Michie [13] (see Table 2 for a list of techniques). Where the behavior change techniques used by the experimental intervention were the same as those in the comparison intervention (eg, [35]), the experimental intervention was coded as not using any behavior change techniques.

Mode of Delivery

Mode of delivery was coded using a novel coding scheme developed by the present authors. For convenience, we divided mode of delivery into (i) automated functions, (ii) communicative functions, and (iii) use of supplementary modes. Each category included a list of delivery modes, and we marked whether or not each intervention used that mode. Automated functions included: (a) the use of an enriched information environment (eg, supplementary content and links, testimonials, videos, or games), (b) automated tailored feedback based on individual progress monitoring (eg, comparison to norms or goals, reinforcing messages, or coping messages), and (c) automated follow-up messages (eg, reminders, tips, newsletters, encouragement). Communicative functions included: (d) access to an advisor to request advice (eg, "ask the expert" facility, expert-led discussion board, or chat sessions), (e) scheduled contact with advisor (eg, emails), and (f) peer-to-peer access (eg, buddy systems, peer-to-peer discussions boards, forums, or live chat). Finally, use of supplementary modes included the use of (g) email, (h) telephone, (i) Short Messaging Service (SMS), (j) CD-ROM, or (k) videoconferencing.

The features of intervention delivery that we coded were, to a large extent, constrained by the features that authors typically report and that can be easily and objectively verified (eg, whether text messages were used). The list is not intended to be exhaustive and we recognize that there are other features that may be important but that are not routinely used or reported, or that are hard to measure. For example, navigational format (eg, the extent to which users are "tunnelled" to particular information vs given free choice [54]), entertainment value (eg, use of quizzes, stories, graphics), appearance (eg, color, layout, screen size [18]), and credibility (eg, the extent to which the website cites sources, credentials). As Internet-based interventions become more common and standards of reporting improve, it should be relatively easy to integrate these additional delivery features into the present coding scheme.

Meta-analytic Strategy

We used Hedges *g* as the primary estimate of effect size for each intervention. Hedges *g* is the difference between the two means (for experimental and control conditions, respectively) divided by the pooled standard deviation. Computations were undertaken using Comprehensive Meta-Analysis Version 2 (Biostat, Englewood, NJ, USA) [55] with the exception of meta-regression computations for which we used the weighted least squares regression command in SPSS 15 for Windows (SPSS Inc, Chicago, IL, USA). Weighted average effect sizes (d_+) were based on a random effects model because studies were likely to be "different from one another in ways too complex to capture by a few simple study characteristics" [56]. Effect sizes were interpreted using Cohen's [57] guidelines. According to Cohen, d_+= .20 should be considered a "small" effect size, d_+= .50 is a "medium" effect size, whereas d_+= .80 is a "large" effect size. The homogeneity Q statistic [58] was used to evaluate variability across effect sizes from the primary studies. When Q is statistically significant it indicates that the effect sizes are heterogeneous. For the meta-regressions, ß is beta weight or coefficient assigned to the predictor and *t* (and the associated *P* -value) tests whether the beta weight is significantly different from zero.

Results

Effect of Internet-based Interventions on Health-related Behavior

The weighted average effect size across all interventions was d_+= 0.16 with a 95% confidence interval from 0.09 to 0.23 based on 85 studies (k = 85) and a total of 43,236 participants (see Table 1). This means that the Internet-based interventions had, on average, a small effect on health behavior according to Cohen's criteria [57]. While these qualitative indices are useful for interpreting the findings of systematic reviews, however, statistical effectiveness is not necessarily the same as clinical effectiveness. For example, a relatively small effect of an Internet-based intervention on smoking abstinence could have substantial clinical significance [59]. On the other hand, an Internet-based intervention that produces a reliable change in fat intake has the potential to benefit a larger proportion of the population than an intervention targeted at smokers. Given that much of the cost associated with Internet-based interventions is likely to be incurred at the design and development stage rather than in delivering individual treatments, small effects

with the potential to have an impact on large numbers of people may thus be significant for patient or population health.

We also calculated effect sizes separately for commonly targeted behaviors (see Table 1). Small, but significant, effects on behavior were observed for Internet-based interventions that targeted only physical activity (d_+= 0.24, k = 20, 95% CI 0.09 to 0.38), dietary behavior (d_+= 0.20, k = 10, 95% CI 0.02 to 0.37), or alcohol consumption (d_+= 0.14, k = 9, 95% CI 0.00 to 0.27). Interventions that targeted smoking abstinence tended to have slightly smaller effects on behavior that did not reach statistical significance (d_+= 0.07, k = 12, 95% CI -0.04 to 0.18). Finally, we calculated effect sizes separately for interventions that targeted multiple behaviors (eg, Williamson et al's intervention [50] targeted physical activity and dietary behavior) and those that targeted a single behavior. Interventions that targeted multiple behaviors tended to have slightly smaller effects on behavior (d_+= 0.12, k = 10, 95% CI 0.08 to 0.17) than did interventions that targeted a single behavior (d_+= 0.17, k = 75, 95% CI 0.09 to 0.24), although both effects were statistically significant.

Table 1. Weighted effect sizes (d+) for behavior change as a function of Internet-based interventions by behavior type

Behavior	*k* [a]	*Q* [b]	95% CI	d_+ [c]
Physical activity	20	128.76[f]	0.09-0.38	0.24[e]
Dietary behavior	10	30.82[e]	0.02-0.37	0.20[f]
Alcohol consumption	9	47.45[f]	0.00-0.27	0.14[d]
Smoking abstinence	12	45.46[e]	-0.04 to 0.18	0.07
Interventions targeting multiple behaviors	10	7.90	0.08-0.17	0.12[f]
Interventions targeting a single behavior	75	879.81[f]	0.09-0.24	0.17[f]
All studies	85	896.67[f]	0.09-0.23	0.16[f]

[a]k = the number of interventions included in the estimate of effect size

[b]Q = homogeneity for the subgroup of interventions

[c]d+ = weighted average effect size

[d]*P* < .05

[e]*P* < .01

[f]*P* < .001

Intervention Characteristics

Across all interventions, the homogeneity Q statistic was highly significant (Q = 896.67, *P* < .001), which indicates considerable variability across effect sizes from the primary studies. To examine the impact of intervention characteristics on effect size, we computed the weighted average effect size for behavior change as a function of the theoretical basis of the interventions, the different ways that the interventions used theory, the behavior change techniques, and the mode of delivery. The findings from these analyses are shown in Table 2. Multimedia Appendix 2 shows the characteristics of each intervention.

Use of Theory and Predictors

Of the different uses of theory proposed by Michie and Prestwich's coding scheme [14], theory or predictors were most commonly used to select or develop intervention techniques (k = 37). Over 20% of the interventions, however, mentioned theory (k = 30), linked at least one intervention technique to theory (k = 19), linked at least one theory-relevant construct to an intervention technique (k = 18), or mentioned a target construct as a predictor of behavior (k = 18). Interventions that used theory or predictors to select recipients for the intervention tended to have the largest effects on behavior (d_+= 0.33, k = 3, 95% CI 0.15 to 0.52) with most other uses of theory tending to have smaller effects (Median d_+= 0.19). Overall, meta-regression indicated that increased use of theory had a significant positive impact on effect sizes (ß = 0.22, *t* = 2.00, *P* = .049). Interventions that made extensive use of theory tended to have larger effects on behavior than did interventions that made less extensive or no use of theory.

Theoretical Basis

Only three theories were used by three or more studies to develop the intervention; social cognitive theory (SCT) [22], the transtheoretical model (TTM) [60], and the theory of reasoned action/planned behavior (TPB) [20,61]. Effect sizes associated with interventions based on the TPB tended to have larger effects on behavior (d_+= 0.36, k = 9, 95% CI 0.15 to 0.56) than did interventions based on the TTM (d_+= 0.20, k = 12, 95% CI 0.08 to 0.33) that, in turn, had larger effects than did interventions based on SCT (d_+= 0.15, k = 12, 95% CI 0.04 to 0.25).

Behavior Change Techniques

The most commonly used behavior change techniques (used by 30% or more of interventions) were providing information on the consequences of behavior in general (k = 29), prompting self-monitoring of behavior (k = 28), and identifying barriers and/or problem solving (k = 26). The largest effects on behavior were observed for interventions that provided stress management (d_+= 0.50, 95% CI 0.27 to 0.72) or general communication skills

training (d_+= 0.49, 95% CI 0.25 to 0.73), although these were used by relatively few interventions (k = 5 and 3, respectively). Modeling, relapse prevention/coping planning, facilitating social comparison, goal setting, action planning, and provision of feedback on performance all had effects on behavior that exceeded d_+= 0.20 (Median d_+= 0.28). Finally, a few strategies had small and non-significant effects on behavior: use of follow-up prompts, self-monitoring of behavioral outcome, emotional control training, and provision of information about others approval. Overall, meta-regression indicated that the number of behavior change techniques employed had a significant positive impact on effect size (ß = 0.36, *t* = 3.48, *P* < .001). Interventions that used more techniques tended to have larger effects on behavior than did interventions that used fewer techniques.

Mode of Delivery

Only one mode of delivery was used by 30% or more of interventions–providing an enriched information environment (k = 30). Over 20% of interventions, however, provided access to an advisor to request advice (k = 23), used peer-to-peer access (k = 20), used email in addition to the Internet-based intervention (k = 19), or provided automated tailored feedback (k = 18). For convenience of interpretation, effect sizes for modes of delivery were divided into three subgroups: automated functions, communicative functions, and use of supplementary modes. In terms of automated functions, small, but significant, effects on behavior were observed for interventions that provided automated tailored feedback (d_+= 0.18, k = 18, 95% CI 0.07 to 0.28) or an enriched information environment (d_+= 0.15, k = 30, 95% CI 0.07 to 0.23). Interventions that provided automated follow-up messages tended not to have significant effects on behavior (d_+= 0.09, k = 14, 95% CI -0.01 to 0.19). Of the communicative functions, interventions that provided access to an advisor to request advice tended to have small-to-medium effects on behavior (d_+= 0.29, k = 23, 95% CI 0.16 to 0.42), while smaller effects on behavior were observed for interventions that provided scheduled contact with an advisor (d_+= 0.22, k = 13, 95% CI 0.09 to 0.36) or peer-to-peer access (d_+= 0.20, k = 20, 95% CI 0.09 to 0.21). Finally, use of additional modes appeared to have distinct effects on behavior change with Internet-based interventions that also used text messages having large effects on behavior (d_+= 0.81, k = 4, 95% CI 0.14 to 1.49), Internet-based interventions using the telephone having small-to-medium effects (d_+= 0.35, k = 7, 95% CI 0.09 to 0.61), and interventions using email as an additional mode of delivery tending to have small effects on behavior (d_+= 0.18, k = 19, 95% CI 0.07 to 0.29).

Table 2. Effect sizes[a] by theoretical basis, use of theory, behavior change techniques, and mode of delivery. The numbering for use of theory, behaviour change techniques, and the letters for mode of delivery correspond with those items in the coding frames and Multimedia Appendix 2.

	K[b]	Q[c]	95% CI	d_+[d]
Theoretical Basis				
Theory of reasoned action/planned behavior (TPB) [20,61]	9	108.44[h]	0.15 to 0.56	0.36[g]
Transtheoretical model (TTM) [60]	12	68.99[h]	0.08 to 0.33	0.20[g]
Social cognitive theory (SCT) [22]	12	18.62	0.04 to 0.25	0.15[g]
Elaboration likelihood model (ELM) [62]	2			
Extended parallel process model (EPPM) [63]	1			
Self-regulation theory (SRT) [64]	1			
Precaution adoption process model (PAPM) [65]	1			
Diffusion of innovations model (DIM) [66]	1			
Health belief model (HBM) [67,68]	1			
Social norms theory (SNT) [69]	1			
Use of Theory				
4. Theory/predictors used to select recipients for the intervention	3	2.84	0.15 to 0.52	0.33[h]
9. Group of techniques are linked to a group of constructs/predictors	6	9.85	0.03 to 0.43	0.23[f]
5. Theory/predictors used to select/develop intervention techniques	37	191.40[h]	0.13 to 0.29	0.21[h]
2. Targeted construct mentioned as predictor of behavior	18	60.07[h]	0.11 to 0.31	0.21[g]
6. Theory/predictors used to tailor intervention techniques to recipients	11	67.75[h]	0.07 to 0.34	0.21[g]
1. Theory/model of behavior mentioned	30	161.33[h]	0.11 to 0.28	0.19[h]
8. At least one of the intervention techniques is linked to theory	19	93.65[h]	0.09 to 0.29	0.19[g]
3. Intervention based on single theory	12	57.13[h]	0.05 to 0.32	0.18[f]
10. All theory-relevant constructs are linked to intervention techniques	10	47.70[h]	-0.02 to 0.37	0.18
11. At least one of the theory-relevant constructs is linked to an intervention technique	18	70.63[h]	0.07 to 0.27	0.17[g]
7. All intervention techniques are linked to theory	2			
Behavior Change Technique				
35. Stress management	5	6.73	0.27 to 0.72	0.50[h]
39. General communication skills training	3	4.38	0.25 to 0.73	0.49[h]
21. Model/demonstrate the behavior	5	24.80[h]	-0.01 to 0.70	0.35[e]
34. Relapse prevention/coping planning	14	38.31[h]	0.17 to 0.47	0.32[h]
27. Facilitate social comparison	4	3.25	0.04 to 0.55	0.29[f]
5. Goal setting (behavior)	25	126.24[h]	0.16 to 0.38	0.27[h]
7. Action planning	18	101.67[h]	0.13 to 0.37	0.25[h]
19. Provide feedback on performance	19	77.38[h]	0.09 to 0.34	0.22[g]
8. Barrier identification/problem solving	26	112.52[h]	0.10 to 0.30	0.20[h]

	K[b]	Q[c]	95% CI	d_+[d]
20. Provide instruction	25	97.95[h]	0.13 to 0.28	0.20[h]
22. Teach to use prompts/cues	3	5.45	-0.17 to 0.57	0.20
4. Provide normative information about others' behavior	16	94.32[h]	0.07 to 0.28	0.18[g]
28. Plan social support/social change	15	41.32[h]	0.10 to 0.27	0.18[h]
13. Provide rewards for behavior	7	7.17	0.09 to 0.28	0.18[h]
16. Prompt self-monitoring of behavior	28	80.81[h]	0.07 to 0.24	0.16[h]
1. Provide information on the consequences in general	29	114.14[h]	0.06 to 0.21	0.14[h]
2. Provide information on the consequences for individual	12	47.57[h]	0.04 to 0.24	0.14[g]
26. Use of follow up prompts	5	39.35[h]	-0.10 to 0.35	0.13
17. Prompt self-monitoring of behavioral outcome	13	45.73[h]	-0.03 to 0.26	0.12
12. Reinforcing effort toward behavior	3	2.89	0.02 to 0.19	0.11[f]
36. Emotional control training	11	35.39[h]	-0.03 to 0.22	0.09
3. Provide information about others' approval	5	10.48[f]	-0.11 to 0.23	0.06
6. Goal setting (outcome)	2			
10. Prompt review of behavioral goals	2			
14. Shaping	2			
23. Environmental restructuring	2			
25. Prompt practice	2			
24. Agree behavioral contract	1			
31. Fear Arousal	1			
32. Prompt self-talk	1			
37. Motivational interviewing	1			
9. Set graded tasks	0			
11. Prompt review of outcome goals	0			
15. Prompting generalisation of behavior	0			
18. Prompting focus on past success	0			
29. Prompt identification as role model	0			
30. Prompt anticipated regret	0			
33. Prompt use of imagery				
38. Time management				
40. Provide non-specific social support				
Mode of Delivery: Automated Functions				
b. Automated tailored feedback	18	83.75[h]	0.07 to 0.28	0.18[g]
a. Enriched information environment	30	117.24[h]	0.07 to 0.23	0.15[h]
c. Automated follow-up messages	14	49.81[h]	-0.01 to 0.19	0.09
Mode of Delivery: Communicative Functions				
d. Access to advisor to request advice	23	121.15[h]	0.16 to 0.42	0.29[h]
e. Scheduled contact with advisor	13	35.70[h]	0.09 to 0.36	0.22[g]

	K[b]	Q[c]	95% CI	d_+[d]
f. Peer-to-peer access	20	88.21[h]	0.09 to 0.21	0.20[h]
Mode of Delivery: Additional Modes				
i. Text message (SMS)	4	39.22[h]	0.14 to 1.49	0.81[a]
h. Telephone	7	19.02[g]	0.09 to 0.61	0.35[g]
g. Email	19	143.98[h]	0.07 to 0.29	0.18[g]
j. CD-ROM	1			
k. Videoconferencing	1			

[a]Effect sizes are ordered within category by size of effect. Characteristics supported by less than three interventions were not examined in order to ensure reliable evaluations of the impact of particular intervention characteristics on effect size.

[b]k = the number of interventions included in the estimate of effect size

[c]Q = homogeneity across the subgroup of interventions

[d]d+ = weighted average effect size

[e]Removing Mikolajczak et al [70] from the evaluation of the effects of modeling on behavior change rendered the effect size significant (k = 4, Q = 13.84, 95% CI 0.14 to 0.84, d+ = 0.49,*P* = .006)

[f]*P* < .05

[g]*P* < .01

[h]*P* < .001

Discussion

Overall Findings

The primary aim of the present review was to relate the characteristics of Internet-based interventions to their effectiveness in promoting health behavior change. Like previous reviews, the interventions tended to have variable effects on behavior (ie, the homogeneity Q statistic was significant), and the average effect on behavior was statistically small. Thus, while some interventions had very large effects (d > 1.00) on behavior (eg, [21,71,72]), others were found to have small or even negative effects on behavior (eg, [73,74]). The considerable variability in the effectiveness of Internet-based interventions makes it important to systematically identify the characteristics of effective interventions and to relate these to effect size.

Use of Theory

Interventions differed substantially in their use of theory, but more extensive use of theory was associated with larger effect sizes. This finding is consistent with assertions that interventions can benefit from using behavior change theory [23-27] and extends the evidence base to interventions delivered on the Internet. Three theories–social cognitive theory (SCT) [22], the transtheoretical model (TTM) [60], and the theory of reasoned action/planned behavior (TPB) [20,61]–were used much more frequently than others. However, only the use of the TPB to inform intervention design led to substantially larger effects than were observed across all interventions. Effect sizes were small-to-medium, comparable to those reported in reviews of non-Internet interventions that used the TPB to develop the intervention [75], and were not simply the consequence of TPB interventions targeting a different set of behaviors. (Interventions based on the TPB targeted a similar range of health-related behaviors to those based on the TTM or SCT.) The observed effectiveness of the TPB in promoting health behavior change stands in contrast to recent assertions that the TPB is primarily a predictive model rather than a model of behavior change that can inform interventions (eg, [76]). However, the heterogeneity of effects across findings means that the findings should be treated with caution and should provide an empirical basis for experimental studies that can demonstrate cause and effect [77,78]. Such studies are also important because Michie and Prestwich's coding of use of theory [14] used in the present review is, necessarily, based on what is reported in the manuscripts; it is of course possible that manuscripts can report having used theory without actually having done so (and vice versa).

Behavior Change Techniques

The finding that interventions that incorporated more behavior change techniques tended to have larger effects than interventions that incorporated fewer techniques justified the investment in relatively elaborate interventions. This finding may be a consequence of different techniques targeting different aspects of the behavior change process [18], and future research might usefully consider how particular combinations of techniques might be especially effective in promoting behavior change [33]. However, there is also evidence that very simple interventions can prove effective in some contexts (eg, providing instruction for influencing online food purchases [79] and if-then planning for promoting dental flossing [7]), and issues of cost versus benefit should always be a consideration in designing interventions to promote health behavior change [80]. Tate et al [81] provide a useful discussion of cost versus effectiveness in relation to Internet-based interventions, and we echo their call for future research to collect cost-effectiveness data.

The two behavior change techniques that were associated with the greatest changes in behavior were stress management and general communication skills training. It is interesting that both techniques influence behavior change indirectly via mechanisms such as facilitating problem-solving, promoting self-efficacy [82], or diminishing the impact of stressors that may prevent behavior change [83]. However, relatively few interventions employed these techniques, so the findings should be treated with caution and form the basis for future research. Given the effectiveness of stress management training, it is perhaps surprising that emotional control training was less effective in promoting behavior change. Of the 11 interventions (45%) that incorporated emotional control training, 5 reported negative effect sizes on behavior [31,32,84,85]. Authors reported that in many of these interventions they simply included "strategies to manage mood [85]"or "information on ... dealing with relationships and feelings [31]." In contrast, stress management training tended to be more intensive. For example, the intervention reported by Hänggi [86] incorporated 4 stress management modules that were based on cognitive behavioral principles. Again, these differences might form a useful basis for future empirical investigation.

Two other findings in relation to behavior change techniques warrant comment. First, it was notable that providing information about others' approval (subjective or injunctive norms) seemed to be less effective than providing normative information about others' behavior (descriptive norms, d_+= 0.06 and 0.18, respectively). This finding supports the distinction between the two types of normative influence [87] and research that shows that descriptive norms can exert a more powerful effect on behavior and decision making than injunctive norms (eg, [88,89]). Second, effect sizes associated with modeling, while substantial overall, were also highly variable rendering the overall estimate of effectiveness non-significant. Modeling is usually used to boost self-efficacy [22], and the present interventions tended to incorporate embedded videos demonstrating the focal behavior within the online intervention (eg, [70,90,91]). The variability in effect sizes in the present review was primarily caused by Mikolajczak et al's "Queermasters" intervention [70], which reported a negative effect on uptake of HIV testing at the three month follow up (d = -0.23). The authors attributed this finding to the relatively short follow-up, which may not have given participants opportunity to act on their newly formed positive intentions. Removing Mikolajczak et al from the evaluation of the effects of modeling on behavior change rendered the effect size significant (k = 4, Q = 13.84, 95% CI = 0.14-0.84, d_+= 0.49, P = .006).

Mode of Delivery

The present review developed a novel coding scheme for the mode by which Internet-based interventions are delivered. Dividing mode of delivery into automated functions, communicative functions, and use of supplementary modes proved informative, with distinct effects being identified within each category. Text messages were highly effective and used in several ways: to promote interaction with the intervention [4,5], send motivational messages (eg, reminders of the benefits of exercise [37]), challenge dysfunctional beliefs [71], or provide a cue to action [35]. Use of communicative functions, especially access to an advisor to request advice, also tended to be effective. It may be that, although the Internet provides a suitable medium for delivering interventions, personal contact via email [92], online [93,94], or text message [95] helps to support behavior change.

Conclusion

The present review is, to our knowledge, the first to systematically code the characteristics of Internet-based interventions designed to promote behavior change and to link these characteristics to effect size. The strengths of the review are the systematic, meta-analytic approach, the use of established coding frames where possible, and the large number of different interventions that focus on a range of different behaviors. The findings suggest that the effectiveness of Internet-based interventions is associated with more extensive use of theory (in particular the use of the theory of planned behavior), inclusion of more behavior change techniques, and use of additional methods of interacting with participants (especially text messages). The review provides a framework for research that can contribute to a science of Internet-based interventions [1] and our findings provide a rationale for investing in more intensive theory-based interventions that incorporate multiple behavior change techniques and modes of delivery. However, the heterogeneity of effects across findings and the relatively small number of interventions associated with some characteristics mean that the findings should be treated with caution and provide an empirical basis for experimental studies that can demonstrate cause and effect.

Acknowledgments

This review was inspired by a workshop on Internet-based behaviour change interventions in addiction sponsored by the Society for the Study of Addiction. The authors would like to thank Hongmei Han, Marney White, and Donald Williamson for providing additional information concerning their research. We also thank Craig Whittington for statistical assistance and Robert West for helpful comments on earlier drafts of this manuscript. This review was funded in part by an ESRC grant (RES-149-25-1069) awarded to LY and SM. This grant funds the Southampton "LifeGuide" node of the National Centre for e-Social Science (www.lifeguideonline.org) and supported JJ.

Multimedia Appendix 1

Effect Sizes for Interventions Included in the Meta-Analysis

[PDF file (Adobe PDF),60 KB - jmir_v12i1e4_app1.pdf]

Multimedia Appendix 2

Intervention Characteristics for Interventions Included in the Meta-Analysis

[PDF file (Adobe PDF),68 KB - jmir_v12i1e4_app2.pdf]

References

1. Ritterband LM, Tate DF. The science of internet interventions. Introduction. Ann Behav Med 2009 Aug;38(1):1-3. [Medline: 19816750] [doi: 10.1007/s12160-009-9132-5]
2. Internet World Stats. World Internet Users and Population Stats URL: http://www.internetworldstats.com/stats.htm [accessed 2009 Oct 6] [WebCite Cache ID 5kK6ed847]
3. Murray E, Khadjesari Z, White IR, Kalaitzaki E, Godfrey C, McCambridge J, et al. Methodological challenges in online trials. J Med Internet Res 2009 Apr;11(2):e9 [FREE Full text] [Medline: 19403465] [doi: 10.2196/jmir.1052]
4. Brendryen H, Kraft P. Happy ending: a randomized controlled trial of a digital multi-media smoking cessation intervention. Addiction 2008 Mar;103(3):478-484. [Medline: 18269367] [doi: 10.1111/j.1360-0443.2007.02119.x]
5. Brendryen H, Drozd F, Kraft P. A digital smoking cessation program delivered through internet and cell phone without nicotine replacement (happy ending): randomized controlled trial. J Med Internet Res 2008;10(5):e51 [FREE Full text] [Medline: 19087949] [doi: 10.2196/jmir.1005]
6. Skår S, Sniehotta FF, Molloy GJ, Prestwich A, Araújo-Soares V. Do brief online planning interventions increase physical activity amongst university students? A randomised controlled trial. Psychol Health. In press .
7. Sniehotta FF, Araújo Soares V, Dombrowski SU. Randomized controlled trial of a one-minute intervention changing oral self-care behavior. J Dent Res 2007 Jul;86(7):641-645. [Medline: 17586711] [doi: 10.1177/154405910708600711]
8. Portnoy DB, Lori AJ, Sheldon S, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction a meta-analysis of 75 randomized controlled trials, 1988–2007. Prev Med 2008 Jul;47(1):3-16. [Medline: 18403003]
9. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [Medline: 19832783] [doi: 10.1111/j.1360-0443.2009.02710.x]
10. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007 Jul;33(1):54-64. [Medline: 17572313] [doi: 10.1016/j.amepre.2007.02.041]
11. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006 Feb;31(2):264-277. [Medline: 15950392] [doi: 10.1016/j.addbeh.2005.05.002]
12. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [Medline: 15631964] [doi: 10.2196/jmir.6.4.e40]
13. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [Medline: 18624603] [doi: 10.1037/0278-6133.27.3.379]
14. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychol 2010 Jan;29(1):1-8. [Medline: 20063930] [doi: 10.1037/a0016939]
15. Bootzin R. Behavior Modification and Therapy: An Introduction. Cambridge, MA: Winthrop; 1975.
16. Kanfer FHE, Goldstein APE. Helping People Change: A Textbook of Methods. New York, NY: Pergamon Press; 1986.
17. Hardeman W, Griffin S, Johnston M, Kinmonth AL, Wareham NJ. Interventions to prevent weight gain: a systematic review of psychological models and behaviour change methods. Int J Obes Relat Metab Disord 2000 Feb;24(2):131-143. [Medline: 10702762] [doi: 10.1038/sj.ijo.0801100]
18. 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. [Medline: 19802647] [doi: 10.1007/s12160-009-9133-4]
19. Spittaels H, De Bourdeaudhuij I, Brug J, Vandelanotte C. Effectiveness of an online computer-tailored physical activity intervention in a real-life setting. Health Educ Res 2007 Jun;22(3):385-396 [FREE Full text] [Medline: 16971674] [doi: 10.1093/her/cyl096]
20. Ajzen I. The Theory of Planned Behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
21. Carr LJ, Bartee RT, Dorozynski C, Broomfield JF, Smith ML, Smith DT. Internet-delivered behavior change program increases physical activity and improves cardiometabolic disease risk factors in sedentary adults: results of a randomized controlled trial. Prev Med 2008 May;46(5):431-438. [Medline: 18207228] [doi: 10.1016/j.ypmed.2007.12.005]
22. Bandura A. Human agency in social cognitive theory. Am Psychol 1989 Sep;44(9):1175-1184. [Medline: 2782727] [doi: 10.1037/0003-066X.44.9.1175]
23. Griffin S, Kinmonth AL, Skinner C, Kelly J. Educational and Psychosocial Interventions for Adults with Diabetes. London, United Kingdom: British Diabetic Association; 1999.

24. Marteau T, Dieppe P, Foy R, Kinmonth AL, Schneiderman N. Behavioural medicine: changing our behaviour. BMJ 2006 Feb 25;332(7539):437-438 [FREE Full text] [Medline: 16497737] [doi: 10.1136/bmj.332.7539.437]
25. Michie S, Sheeran P, Rothman A. Advancing the science of behaviour change. Invited editorial. Psychol Health 2007 Apr;22(3):249-253. [doi: 10.1080/14768320701233582]
26. Painter JE, Borba CPC, Hynes M, Mays D, Glanz K. The use of theory in health behavior research from 2000 to 2005: a systematic review. Ann Behav Med 2008 Jun;35(3):358-362. [Medline: 18633685] [doi: 10.1007/s12160-008-9042-y]
27. Rothman AJ. "Is there nothing more practical than a good theory?": Why innovations and advances in health behavior change will arise if interventions are used to test and refine theory. Int J Behav Nutr Phys Act 2004 Jul 27;1(1):11 [FREE Full text] [Medline: 15279674] [doi: 10.1186/1479-5868-1-11]
28. Jeffery RW. How can Health Behavior Theory be made more useful for intervention research? Int J Behav Nutr Phys Act 2004 Jul 23;1(1):10 [FREE Full text] [Medline: 15272938] [doi: 10.1186/1479-5868-1-10]
29. Oxman AD, Fretheim A, Flottorp S. The OFF theory of research utilization. J Clin Epidemiol 2005 Feb;58(2):113-6; discussion 117-20. [Medline: 15680741] [doi: 10.1016/j.jclinepi.2004.10.002]
30. Albarracín D, Gillette JC, Earl AN, Glasman LR, Durantini MR, Ho MH. A test of major assumptions about behavior change: a comprehensive look at the effects of passive and active HIV-prevention interventions since the beginning of the epidemic. Psychol Bull 2005 Nov;131(6):856-897 [FREE Full text] [Medline: 16351327] [doi: 10.1037/0033-2909.131.6.856]
31. Patten CA, Croghan IT, Meis TM, Decker PA, Pingree S, Colligan RC, et al. Randomized clinical trial of an Internet-based versus brief office intervention for adolescent smoking cessation. Patient Educ Couns 2006 Dec;64(1-3):249-258. [Medline: 16616449] [doi: 10.1016/j.pec.2006.03.001]
32. Muñoz RF, Lenert LL, Delucchi K, Stoddard J, Perez JE, Penilla C, et al. Toward evidence-based Internet interventions: A Spanish/English Web site for international smoking cessation trials. Nicotine Tob Res 2006 Feb;8(1):77-87. [Medline: 16497602] [doi: 10.1080/14622200500431940]
33. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [Medline: 19916637] [doi: 10.1037/a0016136]
34. Marks JT, Campbell MK, Ward DS, Ribisl KM, Wildemuth BM, Symons MJ. A comparison of Web and print media for physical activity promotion among adolescent girls. J Adolesc Health 2006 Jul;39(1):96-104. [Medline: 16781967] [doi: 10.1016/j.jadohealth.2005.11.002]
35. Moore MJ, Soderquist J, Werch C. Feasibility and efficacy of a binge drinking prevention intervention for college students delivered via the Internet versus postal mail. J Am Coll Health 2005;54(1):38-44. [Medline: 16050327] [doi: 10.3200/JACH.54.1.38-44]
36. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008 Aug;10(3):e22 [FREE Full text] [Medline: 18725354] [doi: 10.2196/jmir.1030]
37. Hurling R, Fairley BW, Dias MB. Internet-based exercise intervention systems: Are more interactive designs better? Psychol Health 2006;21(6):757-772. [doi: 10.1080/14768320600603257]
38. Webb TL, Sheeran P. Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychol Bull 2006 Mar;132(2):249-268. [Medline: 16536643] [doi: 10.1037/0033-2909.132.2.249]
39. 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 Feb 15;4(1):e3 [FREE Full text] [Medline: 11956035] [doi: 10.2196/jmir.4.1.e3]
40. Couper MP, Peytchev A, Strecher VJ, Rothert K, Anderson J. Following up nonrespondents to an online weight management intervention: randomized trial comparing mail versus telephone. J Med Internet Res 2007 Jun;9(2):e16 [FREE Full text] [Medline: 17567564] [doi: 10.2196/jmir.9.2.e16]
41. Murray E, McCambridge J, Khadjesari Z, White IR, Thompson SG, Godfrey C, et al. The DYD-RCT protocol: an on-line randomised controlled trial of an interactive computer-based intervention compared with a standard information website to reduce alcohol consumption among hazardous drinkers. BMC Public Health 2007;7(1):306 [FREE Full text] [Medline: 17963483] [doi: 10.1186/1471-2458-7-306]
42. 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 May;9(2):e10 [FREE Full text] [Medline: 17513281] [doi: 10.2196/jmir.9.2.e10]
43. Dijkstra A, De Vries H, Roijackers J. Long-term effectiveness of computer-generated tailored feedback in smoking cessation. Health Educ Res 1998 Jun;13(2):207-214 [FREE Full text] [Medline: 10181019] [doi: 10.1093/her/13.2.207]
44. Strecher VJ, Shiffman S, West R. Moderators and mediators of a web-based computer-tailored smoking cessation program among nicotine patch users. Nicotine Tob Res 2006 Dec;8(Suppl 1):S95-101. [Medline: 17491176] [doi: 10.1080/14622200601039444]
45. McConnon A, Kirk SF, Cockroft JE, Harvey EL, Greenwood DC, Thomas JD, et al. The Internet for weight control in an obese sample: results of a randomised controlled trial. BMC Health Serv Res 2007;7(1):206 [FREE Full text] [Medline: 18093289] [doi: 10.1186/1472-6963-7-206]

46. An LC, Perry CL, Lein EB, Klatt C, Farley DM, Bliss RL, et al. Strategies for increasing adherence to an online smoking cessation intervention for college students. Nicotine Tob Res 2006 Dec;8(Suppl 1):S7-12. [Medline: 17491165] [doi: 10.1080/14622200601039881]
47. Hänggi Y. Kann web-basierte Stressprävention erfolgreich sein? Erfahrungen mit dem Online-Elterntraining gegen Familienstress. Z Klin Psychol 2006;35(3):169-177. [doi: 10.1026/1616-3443.35.3.169]
48. Moher D, Liberati A, Tetzlaff J, Altman DG; Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement (Chinese edition). Zhong Xi Yi Jie He Xue Bao 2009 Sep;7(9):889-896 [FREE Full text] [Medline: 19747449] [doi: 10.3736/jcim20090918]
49. Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions Version 5.0.1. The Cochrane Collaboration 2008.
50. Williamson DA, Walden HM, White MA, York-Crowe E, Newton RL, Alfonso A, et al. Two-year internet-based randomized controlled trial for weight loss in African-American girls. Obesity (Silver Spring) 2006 Jul;14(7):1231-1243. [Medline: 16899804] [doi: 10.1038/oby.2006.140]
51. Hunter JE, Schmidt EL. Methods of Meta-analysis: Correcting Error and Bias in Research Findings. Newbury Park, CA: Sage; 1990.
52. Ashford S, Edmunds J, French DP. What are the most effective techniques for changing self-efficacy to promote physical activity: a meta analysis. 2009 Sep Presented at: DHP Conference; September 2009; Aston, United Kingdom.
53. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39. [Medline: 19806416] [doi: 10.1007/s12160-009-9129-0]
54. Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [Medline: 15914459] [doi: 10.2196/jmir.7.2.e12]
55. Borenstein M, Hedges L, Higgins J, Rothstein H. Comprehensive Meta-analysis Version 2. Englewood, NJ: Biostat; 2005.
56. Cooper H, Hedges L. The Handbook of Research Synthesis. New York, NY: Russell Sage Foundation; 1994:526.
57. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683] [doi: 10.1037/0033-2909.112.1.155]
58. Cochran WG. The combination of estimates from different experiments. Biometrics 1954 Mar;10(1):101-129. [doi: 10.2307/3001666]
59. West R. The clinical significance of "small" effects of smoking cessation treatments. Addiction 2007 Apr;102(4):506-509. [Medline: 17362283] [doi: 10.1111/j.1360-0443.2007.01750.x]
60. Prochaska JO, DiClemente CC. The Transtheoretical Approach: Crossing the Traditional Boundaries of Therapy. Homewood, IL: Dow Jones-Irwin; 1984.
61. Fishbein M, Ajzen I. Belief, Attitude, and Behavior. An Introduction to Theory and Research. Reading, MA: Addison-Wesley; 1975.
62. Petty RE, Cacioppo JT. The Elaboration Likelihood Model of persuasion. In: Berkowitz L, editor. Advances in Experimental Social Psychology (vol. 19). New York, NY: Academic Press; 1986:123-205.
63. Witte K. Putting the fear back into fear appeals: the extended parallel process model. Commun Monogr 1992 Dec;59(4):329-349. [doi: 10.1080/03637759209376276]
64. Kanfer FH. Helping People Change: A Textbook of Methods. 3rd edition. New York, NY: Pergamon Press; 1986.
65. Weinstein ND. The precaution adoption process. Health Psychol 1988;7(4):355-386. [Medline: 3049068] [doi: 10.1037/0278-6133.7.4.355]
66. Rogers EM. Diffusion of Innovations. Glencoe, IL: Free Press; 1962.
67. Rosenstock IM. Why people use health services. Milbank Mem Fund Q 1966 Jul;44(3):Suppl:94-Suppl127. [Medline: 5967464] [doi: 10.2307/3348967]
68. Becker MH. The health belief model and personal health behavior. Health Educ Monog 1974;2:324-473.
69. Perkins HW. The Social Norms Approach to Preventing School and College Age Substance Abuse: A Handbook for Educators, Counsellors, and Clinicians. San Francisco, CA: Jossey-Bass; 2003.
70. Mikolajczak J, Van Breukelen GJP, Kok G, Hospers HJ. Evaluating the Effects of an Online HIV-Prevention Intervention to Promote HIV-Testing Behavior Among Men Who Have Sex With Men (MSM) [dissertation]. Maastricht, Netherlands: Maastricht University; 2008.
71. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007 Apr;9(2):e7 [FREE Full text] [Medline: 17478409] [doi: 10.2196/jmir.9.2.e7]
72. Kim CJ, Kang DH. Utility of a Web-based intervention for individuals with type 2 diabetes: the impact on physical activity levels and glycemic control. Comput Inform Nurs 2006;24(6):337-345. [Medline: 17108753] [doi: 10.1097/00024665-200611000-00008]
73. Frosch DL, Kaplan RM, Felitti VJ. A randomized controlled trial comparing internet and video to facilitate patient education for men considering the prostate specific antigen test. J Gen Intern Med 2003 Oct;18(10):781-787 [FREE Full text] [Medline: 14521639] [doi: 10.1046/j.1525-1497.2003.20911.x]

74. Rodriguez HP, von Glahn T, Rogers WH, Chang H, Fanjiang G, Safran DG. Evaluating patients' experiences with individual physicians: a randomized trial of mail, internet, and interactive voice response telephone administration of surveys. Med Care 2006 Feb;44(2):167-174. [Medline: 16434916] [doi: 10.1097/01.mlr.0000196961.00933.8e]
75. Hardeman W, Johnston M, Johnston DW, Bonetti D, Wareham NJ, Kinmonth AL. Application of the Theory of Planned Behaviour in behaviour change interventions: a systematic review. Psychol Health 2002;17:123-158.
76. Sniehotta FF. Towards a theory of intentional behaviour change: plans, planning, and self-regulation. Br J Health Psychol 2009 May;14(Pt 2):261-273. [Medline: 19102817] [doi: 10.1348/135910708X389042]
77. Sheeran P. Does changing cognitions cause health behaviour change? Psychol Health 2006;21:S137.
78. Sniehotta FF. An experimental test of the Theory of Planned Behaviour. Applied Psychology: Health and Well-Being 2009 May;1(2):257-270. [doi: 10.1111/j.1758-0854.2009.01013.x]
79. Huang A, Barzi F, Huxley R, Denyer G, Rohrlach B, Jayne K, et al. The effects on saturated fat purchases of providing internet shoppers with purchase- specific dietary advice: a randomised trial. PLoS Clin Trials 2006 Sep;1(5):e22 [FREE Full text] [Medline: 17013429] [doi: 10.1371/journal.pctr.0010022]
80. Gold BC, Burke S, Pintauro S, Buzzell P, Harvey-Berino J. Weight loss on the web: A pilot study comparing a structured behavioral intervention to a commercial program. Obesity (Silver Spring) 2007 Jan;15(1):155-164. [Medline: 17228043] [doi: 10.1038/oby.2007.520]
81. Tate DF, Finkelstein EA, Khavjou O, Gustafson A. Cost effectiveness of internet interventions: review and recommendations. Ann Behav Med 2009 Aug;38(1):40-45 [FREE Full text] [Medline: 19834778] [doi: 10.1007/s12160-009-9131-6]
82. Monti PM, Rohsenow DJ, Swift RM, Gulliver SB, Colby SM, Mueller TI, et al. Naltrexone and cue exposure with coping and communication skills training for alcoholics: treatment process and 1-year outcomes. Alcohol Clin Exp Res 2001 Nov;25(11):1634-1647. [Medline: 11707638] [doi: 10.1111/j.1530-0277.2001.tb02170.x]
83. Shiffman S. Coping with temptations to smoke. J Consult Clin Psychol 1984 Apr;52(2):261-267. [Medline: 6715652] [doi: 10.1037/0022-006X.52.2.261]
84. Jones M, Luce KH, Osborne MI, Taylor K, Cunning D, Doyle AC, et al. Randomized, controlled trial of an internet-facilitated intervention for reducing binge eating and overweight in adolescents. Pediatrics 2008 Mar;121(3):453-462 [FREE Full text] [Medline: 18310192] [doi: 10.1542/peds.2007-1173]
85. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. J Med Internet Res 2008 Nov;10(5):e40 [FREE Full text] [Medline: 19017582] [doi: 10.2196/jmir.993]
86. Hänggi Y. Web-based prevention: the online parental training in coping with stress in the family. 2004 Presented at: European Society on Family Relations; September, 2004; Fribourg, Switzerland.
87. Reno RR, Cialdini RB, Kallgren CA. The transsituational influence of social norms. J Pers Soc Psychol 1993 Jan;64(1):104-112. [doi: 10.1037/0022-3514.64.1.104]
88. Conner M, McMillan B. Interaction effects in the theory of planned behaviour: studying cannabis use. Br J Soc Psychol 1999 Jun;38(2):195-222. [Medline: 10392450] [doi: 10.1348/014466699164121]
89. Poliakoff E, Webb TL. What factors predict scientists' intentions to participate in public engagement of science activities? Sci Commun 2007;29(2):242-263. [doi: 10.1177/1075547007308009]
90. Mevissen FEF, Ruiter RAC, Meertens RM, Zimbile FR, Schaalma HP. Justify your love. Testing an online STI risk communication intervention on condom use and STI testing. Psychol Health. Forthcoming .
91. Wade SL, Carey J, Wolfe CR. The efficacy of an online cognitive-behavioral family intervention in improving child behavior and social competence following pediatric brain injury. Rehabil Psychol 2006 Aug;51(3):179-189. [doi: 10.1037/0090-5550.51.3.179]
92. Chan DS, Callahan CW, Hatch-Pigott VB, Lawless A, Proffitt HL, Manning NE, et al. Internet-based home monitoring and education of children with asthma is comparable to ideal office-based care: results of a 1-year asthma in-home monitoring trial. Pediatrics 2007 Mar;119(3):569-578 [FREE Full text] [Medline: 17332210] [doi: 10.1542/peds.2006-1884]
93. Glasgow RE, Boles SM, McKay HG, Feil EG, Barrera M. The D-Net diabetes self-management program: long-term implementation, outcomes, and generalization results. Prev Med 2003 Apr;36(4):410-419. [Medline: 12649049] [doi: 10.1016/S0091-7435(02)00056-7]
94. Gold MR, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in Health and Medicine. Oxford, United Kingdom: Oxford University Press; 1996.
95. Rodgers A, Corbett T, Bramley D, Riddell T, Wills M, Lin RB, et al. Do u smoke after txt? Results of a randomised trial of smoking cessation using mobile phone text messaging. Tob Control 2005 Aug;14(4):255-261 [FREE Full text] [Medline: 16046689] [doi: 10.1136/tc.2005.011577]
96. An LC, Hennrikus DJ, Perry CL, Lein EB, Klatt C, Farley DM, et al. Feasibility of Internet health screening to recruit college students to an online smoking cessation intervention. Nicotine Tob Res 2007 Jan;9(Suppl 1):S11-S18. [Medline: 17365722] [doi: 10.1080/14622200601083418]
97. Bersamin M, Paschall MJ, Fearnow-Kenney M, Wyrick D. Effectiveness of a Web-based alcohol-misuse and harm-prevention course among high- and low-risk students. J Am Coll Health 2007;55(4):247-254. [Medline: 17319331] [doi: 10.3200/JACH.55.4.247-254]

98. Bosak K, Yates B, Pozehl B. An internet physical activity intervention to reduce coronary heart disease risk in the metabolic syndrome population. Circulation 2007;116(9):319. [doi: 10.1161/CIRCULATIONAHA.107.186213]
99. Buller DB, Woodall WG, Zimmerman DE, Slater MD, Heimendinger J, Waters E, et al. Randomized trial on the 5 a day, the Rio Grande Way Website, a web-based program to improve fruit and vegetable consumption in rural communities. J Health Commun 2008 Apr;13(3):230-249. [Medline: 18569356] [doi: 10.1080/10810730801985285]
100. Celio AA, Winzelberg AJ, Wilfley DE, Eppstein-Herald D, Springer EA, Dev P, et al. Reducing risk factors for eating disorders: comparison of an Internet- and a classroom-delivered psychoeducational program. J Consult Clin Psychol 2000 Aug;68(4):650-657. [Medline: 10965640] [doi: 10.1037/0022-006X.68.4.650]
101. Chiauzzi E, Green TC, Lord S, Thum C, Goldstein M. My student body: a high-risk drinking prevention web site for college students. J Am Coll Health 2005;53(6):263-274 [FREE Full text] [Medline: 15900990] [doi: 10.3200/JACH.53.6.263-274]
102. Christakis DA, Zimmerman FJ, Rivara FP, Ebel B. Improving pediatric prevention via the internet: a randomized, controlled trial. Pediatrics 2006 Sep;118(3):1157-1166 [FREE Full text] [Medline: 16951011] [doi: 10.1542/peds.2006-0209]
103. Christensen H, Leach LS, Barney L, Mackinnon AJ, Griffiths KM. The effect of web based depression interventions on self reported help seeking: randomised controlled trial [ISRCTN77824516]. BMC Psychiatry 2006 Apr;6(1):13 [FREE Full text] [Medline: 16595018] [doi: 10.1186/1471-244X-6-13]
104. Cintron A, Phillips R, Hamel MB. The effect of a web-based, patient-directed intervention on knowledge, discussion, and completion of a health care proxy. J Palliat Med 2006 Dec;9(6):1320-1328. [Medline: 17187540] [doi: 10.1089/jpm.2006.9.1320]
105. Cook RF, Billings DW, Hersch RK, Back AS, Hendrickson A. A field test of a web-based workplace health promotion program to improve dietary practices, reduce stress, and increase physical activity: randomized controlled trial. J Med Internet Res 2007 Jun;9(2):e17 [FREE Full text] [Medline: 17581811] [doi: 10.2196/jmir.9.2.e17]
106. Cussler EC, Teixeira PJ, Going SB, Houtkooper LB, Metcalfe LL, Blew RM, et al. Maintenance of weight loss in overweight middle-aged women through the Internet. Obesity (Silver Spring) 2008 May;16(5):1052-1060. [Medline: 18309301] [doi: 10.1038/oby.2008.19]
107. Etter JF. Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial. J Med Internet Res 2005;7(1):e2 [FREE Full text] [Medline: 15829474] [doi: 10.2196/jmir.7.1.e2]
108. Fordis M, King JE, Ballantyne CM, Jones PH, Schneider KH, Spann SJ, et al. Comparison of the instructional efficacy of Internet-based CME with live interactive CME workshops: a randomized controlled trial. JAMA 2005 Sep 7;294(9):1043-1051 [FREE Full text] [Medline: 16145024] [doi: 10.1001/jama.294.9.1043]
109. Gollings EK, Paxton SJ. Comparison of internet and face-to-face delivery of a group body image and disordered eating intervention for women: a pilot study. Eat Disord 2006;14(1):1-15. [Medline: 16757445] [doi: 10.1080/10640260500403790]
110. Hager R, Hardy A, Aldana S. Evaluation of an Internet, stage-based physical activity intervention. Med Sci Sports Exerc 2001;33(5):S231. [doi: 10.1097/00005768-200105001-01301]
111. Heinicke BE, Paxton SJ, McLean SA, Wertheim EH. Internet-delivered targeted group intervention for body dissatisfaction and disordered eating in adolescent girls: a randomized controlled trial. J Abnorm Child Psychol 2007 Jun;35(3):379-391. [Medline: 17243014] [doi: 10.1007/s10802-006-9097-9]
112. Jacobi C, Morris L, Beckers C, Bronisch-Holtze J, Winter J, Winzelberg AJ, et al. Maintenance of internet-based prevention: a randomized controlled trial. Int J Eat Disord 2007 Mar;40(2):114-119. [Medline: 17080447] [doi: 10.1002/eat.20344]
113. Jago R, Baranowski T, Baranowski JC, Thompson D, Cullen KW, Watson K, et al. Fit for Life Boy Scout badge: outcome evaluation of a troop and Internet intervention. Prev Med 2006 Mar;42(3):181-187. [Medline: 16458955] [doi: 10.1016/j.ypmed.2005.12.010]
114. Japuntich SJ, Zehner ME, Smith SS, Jorenby DE, Valdez JA, Fiore MC, et al. Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine Tob Res 2006 Dec;8(Suppl 1):S59-S67. [Medline: 17491172] [doi: 10.1080/14622200601047900]
115. Kosma M, Cardinal BJ, McCubbin JA. A pilot study of a web-based physical activity motivational program for adults with physical disabilities. Disabil Rehabil 2005 Dec 15;27(23):1435-1442. [Medline: 16418058] [doi: 10.1080/09638280500242713]
116. Kypri K, McAnally HM. Randomized controlled trial of a web-based primary care intervention for multiple health risk behaviors. Prev Med 2005;41(3-4):761-766. [Medline: 16120456] [doi: 10.1016/j.ypmed.2005.07.010]
117. 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 Aug;7(4):e47 [FREE Full text] [Medline: 16236699] [doi: 10.2196/jmir.7.4.e47]
118. 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. [Medline: 17063127] [doi: 10.1097/01.mlr.0000233678.80203.c1]
119. 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] [doi: 10.1016/S0749-3797(03)00111-9]
120. McKay HG, King D, Eakin EG, Seeley JR, Glasgow RE. The diabetes network internet-based physical activity intervention: a randomized pilot study. Diabetes Care 2001 Aug;24(8):1328-1334 [FREE Full text] [Medline: 11473065] [doi: 10.2337/diacare.24.8.1328]

121. Meigs JB, Cagliero E, Dubey A, Murphy-Sheehy P, Gildesgame C, Chueh H, et al. A controlled trial of web-based diabetes disease management: the MGH diabetes primary care improvement project. Diabetes Care 2003 Mar;26(3):750-757 [FREE Full text] [Medline: 12610033] [doi: 10.2337/diacare.26.3.750]
122. Napolitano MA, Fotheringham M, Tate D, Sciamanna C, Leslie E, Owen N, et al. Evaluation of an internet-based physical activity intervention: a preliminary investigation. Ann Behav Med 2003 Apr;25(2):92-99. [Medline: 12704010] [doi: 10.1207/S15324796ABM2502_04]
123. Nguyen HQ, Donesky-Cuenco D, Wolpin S, Reinke LF, Benditt JO, Paul SM, et al. Randomized controlled trial of an internet-based versus face-to-face dyspnea self-management program for patients with chronic obstructive pulmonary disease: pilot study. J Med Internet Res 2008 Apr;10(2):e9 [FREE Full text] [Medline: 18417444] [doi: 10.2196/jmir.990]
124. Oenema A, Brug J, Dijkstra A, de Weerdt I, de Vries H. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. Ann Behav Med 2008 Apr;35(2):125-135. [Medline: 18363076] [doi: 10.1007/s12160-008-9023-1]
125. Paschall MJ, Bersamin M, Fearnow-Kenney M, Wyrick D, Currey D. Short-term evaluation of a web-based college alcohol misuse and harm prevention course (College Alc). J Alcohol Drug Educ 2006;50:49-65.
126. Pike KJ, Rabius V, McAlister A, Geiger A. American Cancer Society's QuitLink: randomized trial of Internet assistance. Nicotine Tob Res 2007 Mar;9(3):415-420. [Medline: 17365773] [doi: 10.1080/14622200701188877]
127. Prestwich A. the use of implementation intentions in promoting goal attainment [dissertation]. Leeds, United Kingdon: University of Leeds; 2003.
128. 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. [Medline: 18155287] [doi: 10.1016/j.ypmed.2007.11.007]
129. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [Medline: 18199300] [doi: 10.1111/j.1360-0443.2007.02063.x]
130. Ritterband LM, Cox DJ, Walker LS, Kovatchev B, McKnight L, Patel K, et al. An Internet intervention as adjunctive therapy for pediatric encopresis. J Consult Clin Psychol 2003 Oct;71(5):910-917. [Medline: 14516239] [doi: 10.1037/0022-006X.71.5.910]
131. Roberto AJ, Zimmerman RS, Carlyle KE, Abner EL. A computer-based approach to preventing pregnancy, STD, and HIV in rural adolescents. J Health Commun 2007 Jan;12(1):53-76. [Medline: 17365349] [doi: 10.1080/10810730601096622]
132. Ross SE, Moore LA, Earnest MA, Wittevrongel L, Lin CT. Providing a web-based online medical record with electronic communication capabilities to patients with congestive heart failure: randomized trial. J Med Internet Res 2004 May 14;6(2):e12 [FREE Full text] [Medline: 15249261] [doi: 10.2196/jmir.6.2.e12]
133. Saitz R, Palfai TP, Freedner N, Winter MR, Macdonald A, Lu J, et al. Screening and brief intervention online for college students: the ihealth study. Alcohol Alcohol 2007;42(1):28-36 [FREE Full text] [Medline: 17130139] [doi: 10.1093/alcalc/agl092]
134. Shimazu A, Kawakami N, Irimajiri H, Sakamoto M, Amano S. Effects of web-based psychoeducation on self-efficacy, problem solving behavior, stress responses and job satisfaction among workers: a controlled clinical trial. J Occup Health 2005 Sep;47(5):405-413 [FREE Full text] [Medline: 16230834] [doi: 10.1539/joh.47.405]
135. Spittaels H, De Bourdeaudhuij I, Vandelanotte C. Evaluation of a website-delivered computer-tailored intervention for increasing physical activity in the general population. Prev Med 2007 Mar;44(3):209-217. [Medline: 17197015] [doi: 10.1016/j.ypmed.2006.11.010]
136. Steele R, Mummery WK, Dwyer T. Using the Internet to promote physical activity: a randomized trial of intervention delivery modes. J Phys Act Health 2007 Jul;4(3):245-260. [Medline: 17846455]
137. Strecher VJ, Shiffman S, West R. Randomized controlled trial of a web-based computer-tailored smoking cessation program as a supplement to nicotine patch therapy. Addiction 2005 May;100(5):682-688. [Medline: 15847626] [doi: 10.1111/j.1360-0443.2005.01093.x]
138. Strecher VJ, McClure JB, Alexander GL, Chakraborty B, Nair VN, Konkel JM, et al. Web-based smoking-cessation programs: results of a randomized trial. Am J Prev Med 2008 May;34(5):373-381 [FREE Full text] [Medline: 18407003] [doi: 10.1016/j.amepre.2007.12.024]
139. Swartz LHG, Noell JW, Schroeder SW, Ary DV. A randomised control study of a fully automated internet based smoking cessation programme. Tob Control 2006 Feb;15(1):7-12 [FREE Full text] [Medline: 16436397] [doi: 10.1136/tc.2003.006189]
140. Thombs DL, Olds RS, Osborn CJ, Casseday S, Glavin K, Berkowitz AD. Outcomes of a technology-based social norms intervention to deter alcohol use in freshman residence halls. J Am Coll Health 2007;55(6):325-332. [Medline: 17517544] [doi: 10.3200/JACH.55.6.325-332]
141. Tomnay JE, Pitts MK, Kuo TC, Fairley CK. Does the Internet assist clients to carry out contact tracing? A randomized controlled trial using web-based information. Int J STD AIDS 2006 Jun;17(6):391-394. [Medline: 16734961] [doi: 10.1258/095646206777323391]
142. van den Berg MH, Ronday HK, Peeters AJ, le Cessie S, van der Giesen FJ, Breedveld FC, et al. Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. Arthritis Rheum 2006 Dec 15;55(6):935-945 [FREE Full text] [Medline: 17139640] [doi: 10.1002/art.22339]

143. Vinokur AD, Merion RM, Couper MP, Jones EG, Dong Y. Educational web-based intervention for high school students to increase knowledge and promote positive attitudes toward organ donation. Health Educ Behav 2006 Dec;33(6):773-786. [Medline: 16923836] [doi: 10.1177/1090198106288596]
144. Walters ST, Vader AM, Harris TR. A controlled trial of web-based feedback for heavy drinking college students. Prev Sci 2007 Mar;8(1):83-88. [Medline: 17136461] [doi: 10.1007/s11121-006-0059-9]
145. White MA, Martin PD, Newton RL, Walden HM, York-Crowe EE, Gordon ST, et al. Mediators of weight loss in a family-based intervention presented over the internet. Obes Res 2004 Jul;12(7):1050-1059. [Medline: 15292468] [doi: 10.1038/oby.2004.132]
146. Winett RA, Anderson ES, Wojcik JR, Winett SG, Bowden T. Guide to health: nutrition and physical activity outcomes of a group-randomized trial of an Internet-based intervention in churches. Ann Behav Med 2007 Jun;33(3):251-261. [Medline: 17600452]
147. Zabinski MF, Pung MA, Wilfley DE, Eppstein DL, Winzelberg AJ, Celio A, et al. Reducing risk factors for eating disorders: targeting at-risk women with a computerized psychoeducational program. Int J Eat Disord 2001 May;29(4):401-408. [Medline: 11285577] [doi: 10.1002/eat.1036]
148. Zabinski MF, Wilfley DE, Calfas KJ, Winzelberg AJ, Taylor CB. An interactive psychoeducational intervention for women at risk of developing an eating disorder. J Consult Clin Psychol 2004 Oct;72(5):914-919. [Medline: 15482051] [doi: 10.1037/0022-006X.72.5.914]

Abbreviations

DIM: diffusion of innovations model
ELM: elaboration likelihood model
EPPM: extended parallel process model
HBM: health belief model
PAPM: precaution adoption process model
SCT: social cognitive theory
SMS: short message service
SNT: social norms theory
SRT: self-regulation theory
TPB: theory of reasoned action/planned behavior
TTM: transtheoretical model

Edited by G Eysenbach; submitted 09.10.09; peer-reviewed by P Kraft, R Botelho; comments to author 30.10.09; revised version received 12.11.09; accepted 16.11.09; published 17.02.10.

Please cite as:
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
URL: http://www.jmir.org/2010/1/e4/
doi:10.2196/jmir.1376
PMID:20164043

Short Paper

Utilization Patterns and User Characteristics of an Ad Libitum Internet Weight Loss Program

Martin Binks[1,2], PhD; Trevor van Mierlo[3,4], BA

[1]Binks Behavioral Health, PLLC, Hillsborough, NC, USA

[2]Division of Medical Psychology, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, NC, USA

[3]Evolution Health Systems Inc, Toronto, ON, Canada

[4]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Martin Binks, PhD
Binks Behavioral Health, PLLC
2801 Canter Drive
Hillsborough, NC 27278
USA
Phone: 1 919 732 5524
Fax: 1 919 732 5524
Email: mbinks@binksbehavioralhealth.com

Abstract

Background: The Internet holds promise for the delivery of evidence-based weight loss treatment to underserved populations. However, most studies do not reflect the more naturalistic and common *ad libitum*, or freely at will, use of the Internet. Randomized clinical trials, for example, typically include at least some direct contact with participants and often have restrictive selection criteria. There is a paucity of research examining utilization patterns of online weight loss programs, particularly in the rapidly expanding direct-to-consumer arena.

Objectives: To examine self-reported characteristics (age, body mass index [BMI], gender), behaviors, and Internet site utilization patterns of a sample of users of a direct-to-consumer *ad libitum* Internet weight loss program.

Methods: This study is based on analysis of archival data from the initial 15 weeks of an ongoing, free, evidence-based, direct-to-consumer Internet weight loss program, the Healthy Weight Center, which included standard information about nutrition, fitness, and behavioral strategies; monitoring tools; and moderated support group message boards. Participants encountered the program through self-directed Internet searches and anonymously registered to utilize the site. Self-reported user characteristics and electronically tracked utilization data were extracted from existing program data, compiled, and examined. Pearson correlations were computed to examine the association of program utilization with age and BMI. One-way analysis of variance (ANOVA) was used for gender comparisons.

Results: We examined data from the first 204 adult users of the program who were classified as either overweight (BMI 25 to < 30 kg/m^2) or obese (BMI ≥ 30 kg/m^2). The mean age of participants was 42.0 years (SD 11.7), 81.9% (167/204) were women, and mean BMI was 32.01 kg/m^2 (SD 6.26). The percent of participants who used program tools was as follows: 13.7%, meal planner; 10.8%, nutrition lookup: 17.6%, activity log; 14.2%, journal; and 22.1%, weight tracker. Participants also used the following educational resources: nutrition, 13.2%; fitness, 6.4%; and behavioral, 7.4%. Of the personal self-assessments available through the program, 57.8% of participants assessed personal barriers, and 50.5% assessed relationship with food. Only 7.8% used the support group message boards. No significant associations between site utilization and age, gender, or BMI were found. Reasons for wanting to lose weight were: health, 87%; appearance, 74%; mobility, 44%; doctor recommendation, 23%; and spouse/friend suggested, 12%. The age participants reported first becoming overweight was young adulthood, 31%; late adulthood, 28%; childhood, 22%; adolescence, 17%; and as a toddler, 3%. Self-perceived factors contributing to weight gain were lack of exercise for 70% of participants, emotions for 62%, overeating for 61%, and slow metabolism for 33%.

Conclusions: Internet weight loss programs reach many people who cannot access traditional treatment. However, users appear not to be optimally utilizing key aspects of the weight loss intervention, such as education, monitoring, and support. This study provides insight into the patterns of *ad libitum* use of an online weight loss program across multiple treatment-related domains in a naturalistic Internet environment.

(J Med Internet Res 2010;12(1):e9) doi:10.2196/jmir.1347

KEYWORDS

Weight loss; obesity; Web-based; self-help; Internet; utilization; intervention

Introduction

The Internet holds promise for the delivery of weight loss programs to underserved populations. However, the majority of published Internet weight loss research is from controlled clinical trials [1]. Clinical trial-based eHealth interventions may face challenges to generalizability relating to the inclusion of direct personal contact and extensive screening that is not part of a more naturalistic *ad libitum*, or freely at will, use of the Internet, as is the case with most Internet users' experiences [2,3]. It also has been shown that eHealth programs are susceptible to poor utilization and attrition, which further limits their utility [4].

Internet-based weight loss research typically focuses on treatment outcomes, overall attrition rates, site logins and nutrition/weight tracking without consideration of the degree to which the therapeutic intervention (eg, accessing assessments, educational content, and support groups) was delivered [5]. If we are to develop effective Internet weight loss programs, we must ensure that the core treatment elements are being delivered. Individual tailoring of content through user self-assessments may be an effective approach to improving utilization. By using brief online questionnaires to assess the range of issues that may impact weight loss we may be able to deliver more personally relevant content to each individual user that addresses their particular area(s) of concern. Scientifically rigorous (ie, meta-analytic) review of research evaluating the efficacy of this approach is unavailable; however, narrative reviews have suggested that self-assessments may be too time consuming and thus unacceptable to the user [6,7]. Therefore, consideration of the degree to which participants use interactive self-assessments represents an important step in determining their feasibility.

In this analysis of archival data we examine website utilization data gathered from a free, evidence-based, anonymously utilized, direct-to-consumer weight loss program that includes nutrition, fitness, and behavioral information; self-assessments; monitoring tools; and moderated support group message boards.

Methods

Setting and Program Description

The Healthy Weight Center [8] is a free access, evidence-based [9,10] Internet weight loss program that provides nutrition, fitness, and behavioral information; monitoring tools; interactive assessments with individually tailored feedback; and moderated support group message boards (see Multimedia Appendix 1). The Internet intervention is based on a review of design principles and consumer testing across multiple eHealth platforms (ie, smoking cessation, depression, panic disorder, and problem drinking) and usability testing of this and our other eHealth platforms [11]. The program utilizes free-form matrix design, that is, all program elements are available to every user. The program also provides an online program guide to facilitate full utilization of the program elements.

The program was available for free to anyone having access to the World Wide Web from any geographic region of the world. There was no advertising or promotion of the program. A link to the program was added to the Evolution Health corporate site [12], and potential users could also encounter the program using publicly available Internet searches. The period of data collection for this study was the initial 15-week period of program availability from May to Sept 2008. The program continues to be available to the public. There was no formal recruitment for the program. To enroll, users completed the online registration process consisting of 14 questions that included weight, height, date of birth, gender, email address, occupation and questions relating to weight and dieting history. Registrants were also asked to electronically endorse the program disclosure agreement explaining that data were being collected anonymously and that this unidentifiable information would be used for research purposes.

Subjects

The subjects of this study were the first 204 individuals who registered for the online weight loss program and who met the following inclusion criteria: (1) 18 years or older, (2) body mass index (BMI) classified as overweight (BMI 25 to < 30 kg/m^2) or obese (BMI ≥ 30 kg/m^2); (3) completion of all questions contained in the online registration process; and (4) endorsement of the program disclosure agreement.

Of these 204 participants, 18.1% (37/204) were men, and 81.9% (167/204) were women. The mean age of the participants was 42.0 (SD 11.7) years; 44.4 (SD 12.6) years for men, and 41.5 (SD 11.5) years for women. The mean BMI of the participants was 32.0 (SD 6.3) kg/m^2; 31.2 (SD 5.2) kg/m^2 for men, and 32.2 (SD 6.5) kg/m^2 for women.

Mean weight of the participants was 89.4 (SD 19.9) kg; 99.8 (SD 16.8) kg for men, and 86.9 (SD 19.3) kg for women. All registrants with complete data were included for analysis (N=204).Thirty-three subjects were excluded for missing (BMI) data.

Data Collection

Data for the 204 registrants meeting the inclusion criteria were extracted from the existing Healthy Weight Center database. All online registration questionnaires for the Healthy Weight Center program adhered to international privacy guidelines [13,14]. Procedures were in accordance with the Helsinki Declaration of 1975, as revised in 2008 [15]. Because the study was based on the use of unidentifiable archival data, the study was determined to be exempt from further review.

Personal characteristics and behaviors were based on registrants' self-reported responses to the online questionnaires. We used website analysis tools to determine the extent of program

utilization for each participant. Utilization of each individual program element was defined as accessing that program element at least once: educational information (nutrition, fitness, and behavioral), self-assessments (personal barriers and relationship with food) [16], monitoring tools (weight tracker, exercise tracker, journal, and nutrition lookup), and the moderated support groups.

Statistical Analysis

Personal characteristics/behaviors and utilization data were tabulated as percent reporting. To explore the influence on site utilization of age (eg, younger individuals may be more Internet savvy and thus engage in higher utilization) and BMI (eg, heavier individuals may be more motivated and thus engage in higher utilization), Pearson correlations were computed. One-way analysis of variance (ANOVA) was used to consider gender differences in demographic factors (ie, age, BMI) and program utilization.

Results

Gender

One-way ANOVA revealed no significant differences by gender for age ($F_{1,202}$ = 1.8, P = .18), BMI ($F_{1,202}$ = 0.7, P = .40), and program utilization ($F_{1,202}$ = 0.4, P = .53).

Program Utilization, Age, and BMI

The mean number of program elements utilized was 2.7 (SD 3.9). Pearson correlations revealed no significant correlations of program utilization with age (N = 204; r = -.02; P = .78) or BMI (N = 204; r = -.02; P = .81).

Utilization Patterns

Site utilization data for all users (N= 204) based on accessing each tool at least one time is presented in Table 1. Site tools used for monitoring key weight loss behaviors (meal planner, nutrition lookup, activity log, journal and weight tracker) were not highly utilized. Similarly, educational materials (nutrition, fitness, and behavioral) and support group message boards were not accessed by the majority of users. However, interactive assessments (personal barriers and relationship with food), showed relatively higher utilization.

Personal Characteristics/Behaviors

Data concerning personal characteristics/behaviors were compiled for all participants (N = 204). When participants were asked to indicate all applicable reasons for wanting to lose weight 87% (177/204) endorsed health; 74% (151/204), appearance; 44% (89/204), mobility; 23% (47/204), doctor recommended; and 12% (24/204) indicated that a friend/spouse suggested they lose weight. In terms of the age that users reported first becoming overweight, 31% (63/204) reported young adult onset, 28% (57/204) reported late adult onset, 22% (44/204) reported onset during childhood, 17% (34/204) reported adolescent onset, and 3% (6/204) reported first becoming overweight as a toddler. Of the factors that participants believed contributed most to their weight gain, the most frequently endorsed items were: lack of exercise, endorsed by 70% (143/204); emotions, endorsed by 62% (126/204); overeating, endorsed by 61% (124/204); and slow metabolism, endorsed by 33% (67/204).

Table 1. Utilization of program elements (N=204)

Program Element (see Multimedia Appendix 1)	Utilization n (%)
Meal planner	28 (13.7)
Nutritional data look up tool	22 (10.8)
Activity log	36 (17.6)
Journal	29 (14.2)
Weight tracker	45 (22.1)
Nutrition education	27 (13.2)
Fitness education	13 (6.4)
Behavioral education	15 (7.4)
Support group message board	16 (7.8)
Personal barriers assessment	118 (57.8)
Relationship with food assessment	103 (50.5)

Discussion

Internet weight loss programs that provide evidence-based intervention can reach many individuals who have limited access to traditional treatments. However, website users may not be utilizing Web-based weight loss resources optimally. The standard of care in weight management involves nutrition, fitness, and behavioral education; behavioral self-monitoring; and use of personal support [9,10]. In our study of *ad libitum* use of a Web-based weight loss program, participants did not take full advantage of these essential treatment elements.

In our Internet weight loss program, interactive self-assessments (eg, personal barriers and relationship with food) were more highly utilized than the similarly interactive weight/fitness

tracking tools and journaling. This finding suggests that interactivity alone was not likely responsible for higher utilization. A distinguishing feature of our interactive self-assessments was the promise of personally relevant feedback, which may have made these exercises particularly appealing to the user. Evidence from previous studies suggests that tailoring interventions to individuals may be an effective tool in improving health behavior, but concerns arise about the acceptability to participants of completing the self-assessments that are needed to tailor content [6,7,17]. Our findings suggest that users may actually prefer using interactive self-assessments as compared with other program elements that were available to them in our study (monitoring tools, educational content, support group message boards). This finding highlights the need for architects of Internet-based programs to explore increasing the use of interactive exercises to tailor the user experience in ways that increase personal relevance. This may be preferable to relying on predetermined blocks of educational content to deliver the necessary treatment elements. Furthermore, identifying relevant baseline personal characteristics/behaviors of users may also deserve further consideration. For example, in this study, the majority of users reported that health, mobility, and appearance were important contributors to their decision to engage in the program. Perhaps by tailoring interactive assessments and content to areas of relevance to the individual user, we may be able to approximate the more typical in person health care experience.

Finally, while we did not consider this directly in our study, the literature suggests that utilization of therapeutic program elements may be increased through novel approaches to enhancing the user experience and including incentives such as rewards for completing exercises. Architects of future Internet-based weight loss programs should consider these strategies [18,19,20].

Strengths and Limitations

This study provides insight into how individuals utilize Internet resources in a natural setting as opposed to a clinical trial environment. The relatively fewer barriers to enrollment, based on this program's open availability on the Internet in comparison with the barriers to entry that are typical of Internet-based clinical trials, is a particular strength. Also, few studies have examined the utilization of specific program elements (eg, educational materials, tools, and assessments), focusing instead on broader criteria such as the total number of log-ins. Limitations of our study include the small sample size in comparison with other studies of natural use of Internet health programs and lack of information about users' geographic area (eg, country of residence) each of which may limit generalizability. Another limitation is the lack of information about treatment outcomes and our inability to determine the frequency (multiple accessing of each element) or extent (amount of time spent using each element) with which elements of the program were used, as our data was based on accessing each element at least once. These factors are important considerations for future studies.

Summary

This study provides insight into potential strategies for optimizing program design to improve utilization of core treatment elements of Internet-based weight loss programs. This is a necessary step to ensure quality when widely disseminating weight loss treatments. Future studies should include detailed analysis of utilization patterns and their relationship to both short- and long-term outcomes.

Conflicts of Interest

Dr. Binks has acted as a paid consultant to Evolution Health Systems Inc, Toronto, ON, Canada/San Francisco, CA, USA, which owns the Healthy Weight Center software among other eHealth platforms. Trevor van Mierlo is the CEO of Evolution Health Systems Inc.

Multimedia Appendix 1

Screenshots from the Healthy Weight Center Internet site illustrating the program elements outlined in Table 1

[PPT file (Microsoft Powerpoint), 1380 KB - jmir_v12i1e9_app1.ppt]

References

1. Kirbis-Weinstein PK. A review of weight loss programs delivered via the internet. J Cardiovascular Nursing 2006;2(4):251-258 [FREE Full text]
2. Cunningham JA, van Mierlo T. Methodological issues in the evaluation of Internet-based interventions for problem drinking. Drug Alcohol Rev 2009 Jan;28(1):12-17. [Medline: 19320670]
3. Bonell C, Oakley A, Hargreaves J, Strange V, Rees R. Assessment of generalisability in trials of health interventions: suggested framework and systematic review. BMJ 2006 Aug 12;333(7563):346-349 [FREE Full text] [Medline: 16902217] [doi: 10.1136/bmj.333.7563.346]
4. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar;7(1):e11 [FREE Full text] [Medline: 15829473] [doi: 10.2196/jmir.7.1.e11]
5. Jonasson J, Linné Y, Neovius M, Rössner S. An Internet-based weight loss programme -- a feasibility study with preliminary results from 4209 completers. Scand J Public Health 2009 Jan;37(1):75-82. [Medline: 18684784] [doi: 10.1177/1403494807085066]

6. Neville LM, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. Int J Behav Nutr Phys Act 2009;6(1):30 [FREE Full text] [Medline: 19490649] [doi: 10.1186/1479-5868-6-30]
7. Lustria MLA, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [Medline: 18947966] [doi: 10.1016/j.pec.2008.08.023]
8. Healthy Weight Center. URL: http://www.healthyweightcenter.net/ [accessed 2010-02-01] [WebCite Cache ID 5nECOVnS3]
9. Wing RR, Tate DF. Endotext.com 2002. Behavior modification for obesity. In: Caro JF, editor. Obesity. URL: http://www.endotext.org/obesity/obesity17/obesityframe17.htm [accessed 2010-02-01]
10. ; National Heart, Lung and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services. The Clinical Guidelines on the Identification, Evaluation and Treatment of Overweight and Obesity in Adults: Evidence Report Sep 1998. URL: http://www.ncbi.nlm.nih.gov/bookshelf/br.fcgi?book=obesity [accessed 2010-02-01]
11. van Mierlo T, Cunningham J. Anonymous Internet-based interventions for problem drinkers: design, development and dissemination. Presented at: 35th Annual Alcohol Epidemiology Symposium of the Kettil Bruun Society; June 5, 2009; Copenhagen, Denmark URL: http://www.kbs2009.dk/abstracts-kbs2009.pdf [accessed 2010-02-01] [WebCite Cache ID 5nEDcRziV]
12. Evolution Health Systems Inc. URL: http://www.evolutionhs.com/programs.aspx [accessed 2010-02-13] [WebCite Cache ID 5nW5plLtf]
13. Office of the Privacy Commissioner of Canada. 2000. Updated 2006. Personal Information Protection and Electronic Documents Act URL: http://www.priv.gc.ca/legislation/02_06_01_e.cfm [WebCite Cache ID 5nEGDr6jp]
14. ; U.S. Department of Health & Human Services. Health Insurance Portability and Accountability Act of 1996. p. 104-191 URL: http://www.hhs.gov/ocr/office/index.html [accessed 2010-02-01] [WebCite Cache ID 5Xfm36ydO]
15. ; World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. 7th revision. Presented at: 59th WMA General Assembly; 2008; Seoul, Korea URL: http://infomed.mds-ev.de/sindbad.nsf/ddf5e481ed777582002567cb004313b9/75521880dcc4739dc12574ee00376317/$FILE/DoH_6Rev_SEOUL_2008-10.pdf [accessed 2010-02-13] [WebCite Cache ID 5nWFlChiw]
16. Cappelleri JC, Bushmakin AG, Gerber RA, Leidy NK, Sexton CC, Karlsson J, et al. Evaluating the Power of Food Scale in obese subjects and a general sample of individuals: development and measurement properties. Int J Obes (Lond) 2009 Aug;33(8):913-922. [Medline: 19506564] [doi: 10.1038/ijo.2009.107]
17. de Vries H, Kremers SPJ, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008;22(6):417-425. [Medline: 18677882]
18. Danaher BG, McKay HG, Seeley JR. The information architecture of behavior change websites. J Med Internet Res 2005;7(2):e12 [FREE Full text] [Medline: 15914459] [doi: 10.2196/jmir.7.2.e12]
19. Wing RR, Epstein LH, Marcus M, Shapira B. Strong monetary contingencies for weight loss during treatment and maintenance. Behav Ther 1981;12(5):702-710. [doi: 10.1016/S0005-7894(81)80141-4]
20. Volpp KG, John LK, Troxel AB, Norton L, Fassbender J, Loewenstein G. Financial incentive-based approaches for weight loss: a randomized trial. JAMA 2008 Dec 10;300(22):2631-2637 [FREE Full text] [Medline: 19066383] [doi: 10.1001/jama.2008.804]

Edited by G Eysenbach; submitted 03.09.09; peer-reviewed by M Wanner, J Kabukoba, L Ware; comments to author 27.09.09; revised version received 17.02.10; accepted 21.02.10; published 29.03.10.

Please cite as:
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
URL: http://www.jmir.org/2010/1/e9/
doi:10.2196/jmir.1347
PMID:20350926

Original Paper

Effects of Internet Use on Health and Depression: A Longitudinal Study

Katie Bessière[1], MS; Sarah Pressman[2], PhD; Sara Kiesler[1], PhD; Robert Kraut[1], PhD

[1]Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, United States

[2]Department of Psychology, University of Kansas, Lawrence, United States

Corresponding Author:
Robert Kraut, PhD
Human-Computer Interaction Institute
Carnegie Mellon University
5000 Forbes Ave
Pittsburgh, 15213
United States
Phone: 1 412 268 7693
Fax: 1 412 268 1266
Email: robert.kraut@cmu.edu

Abstract

Background: The rapid expansion of the Internet has increased the ease with which the public can obtain medical information. Most research on the utility of the Internet for health purposes has evaluated the quality of the information itself or examined its impact on clinical populations. Little is known about the consequences of its use by the general population.

Objective: Is use of the Internet by the general population for health purposes associated with a subsequent change in psychological well-being and health? Is the effect different for healthy versus ill individuals? Does the impact of using the Internet for health purposes differ from the impact of other types of Internet use?

Methods: Data come from a national US panel survey of 740 individuals conducted from 2000 to 2002. Across three surveys, respondents described their use of the Internet for different purposes, indicated whether they had any of 13 serious illnesses (or were taking care of someone with a serious illness), and reported their depression. In the initial and final surveys they also reported on their physical health. Lagged dependent variable regression analysis was used to predict changes in depression and general health reported on a later survey from frequency of different types of Internet use at an earlier period, holding constant prior depression and general health, respectively. Statistical interactions tested whether uses of the Internet predicted depression and general health differently for people who initially differed on their general health, chronic illness, and caregiver status.

Results: Health-related Internet use was associated with small but reliable increases in depression (ie, increasing use of the Internet for health purposes from 3 to 5 days per week to once a day was associated with .11 standard deviations more symptoms of depression, P=.002). In contrast, using the Internet for communication with friends and family was associated with small but reliable decreases in depression (ie, increasing use of the Internet for communication with friends and family purposes from 3 to 5 days per week to once a day was associated with .07 standard deviations fewer symptoms of depression, P=.007). There were no significant effects of respondents' initial health status (P=.234) or role as a caregiver (P=.911) on the association between health-related Internet use and depression. Neither type of use was associated with changes in general health (P=.705 for social uses and P=.494 for health uses).

Conclusions: Using the Internet for health purposes was associated with increased depression. The increase may be due to increased rumination, unnecessary alarm, or over-attention to health problems. Additionally, those with unmeasured problems or those more prone to health anxiety may self-select online health resources. In contrast, using the Internet to communicate with friends and family was associated with declines in depression. This finding is comparable to other studies showing that social support is beneficial for well-being and lends support to the idea that the Internet is a way to strengthen and maintain social ties.

(J Med Internet Res 2010;12(1):e6) doi:10.2196/jmir.1149

KEYWORDS

Depression; health; social support; Internet

Edited by G Eysenbach; submitted 07.01.09; peer-reviewed by S Bass; comments to author 05.02.09; revised version received 23.07.09; accepted 30.11.09; published 12.03.10.

Please cite as:
Bessière K, Pressman S, Kiesler S, Kraut R
Effects of Internet Use on Health and Depression: A Longitudinal Study
J Med Internet Res 2010;12(1):e6
URL: http://www.jmir.org/2010/1/e6/
doi:10.2196/jmir.1149
PMID:

Original Paper

Measures of Physical Activity Using Cell Phones: Validation Using Criterion Methods

Christin Bexelius[1], MSc; Marie Löf[1,2], PhD; Sven Sandin[1], MSc; Ylva Trolle Lagerros[3], MD PhD; Elisabet Forsum[2], PhD; Jan-Eric Litton[1], PhD

[1]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

[2]Department of Clinical and Experimental Medicine, Faculty of Health Science, Linköping University, Linköping, Sweden

[3]Clinical Epidemiology Unit, Department of Medicine, Karolinska Institutet, Stockholm, Sweden

Corresponding Author:
Christin Bexelius, MSc
Department of Medical Epidemiology and Biostatistics
Karolinska Institutet
Nobels väg 12A
Stockholm, SE-17177
Sweden
Phone: +46(0) 8 524 87759
Fax: +46(0) 8 314975
Email: Christin.Bexelius@ki.se

Abstract

Background: Physical activity is associated with reduced risks of many chronic diseases. Data collected on physical activity in large epidemiological studies is often based on paper questionnaires. The validity of these questionnaires is debated, and more effective methods are needed.

Objective: This study evaluates repeated measures of physical activity level (PAL) and the feasibility of using a Java-based questionnaire downloaded onto cell phones for collection of such data. The data obtained were compared with reference estimates based on the doubly labeled water method and indirect calorimetry (PAL_{ref}).

Method: Using a Java-based cell phone application, 22 women reported their physical activity based on two short questions answered daily over a 14-day period (PAL_{cell}). Results were compared with reference data obtained from the doubly labeled water method and indirect calorimetry (PAL_{ref}). Results were also compared against physical activity levels assessed by two regular paper questionnaires completed by women at the end of the 14-day period (PAL_{quest1} and PAL_{quest2}). PAL_{cell}, PAL_{quest1}, and PAL_{quest2} were compared with PAL_{ref} using the Bland and Altman procedure.

Results: The mean difference between PAL_{cell} and PAL_{ref} was small (0.014) with narrow limits of agreement (2SD = 0.30). Compared with PAL_{ref}, the mean difference was also small for PAL_{quest1} and PAL_{quest2} (0.004 and 0.07, respectively); however, the limits of agreement were wider (PAL_{quest1}, 2SD = 0.50 and PAL_{quest2}, 2SD = 0.90). The test for trend was statistically significant for PAL_{quest1} (slope of regression line = 0.79, $P = .04$) as well as for PAL_{quest2} (slope of regression line = 1.58, $P < .001$) when compared with PAL_{ref}.

Conclusion: A Java-based physical activity questionnaire administered daily using cell phones produced PAL estimates that agreed well with PAL reference values. Furthermore, the limits of agreement between PAL obtained using cell phones, and reference values were narrower than for corresponding estimates obtained using paper questionnaires. Java-based questionnaires downloaded onto cell phones may be a feasible and cost-effective method of data collection for large-scale prospective studies of physical activity.

(J Med Internet Res 2010;12(1):e2) doi:10.2196/jmir.1298

KEYWORDS

Physical activity, cellular phone, epidemiological methods, data collection

Introduction

Physical activity is associated with reduced risks of several chronic diseases such as cardiovascular diseases, diabetes, and certain cancers [1,2]. To further explore the preventive effects of physical activity, more epidemiological studies with solid data collection are needed. Because energy expenditure affects energy balance and thus body weight and composition, an important health-related consequence of physical activity is the amount of energy expended when being physically active. Energy expenditure in response to physical activity can be measured as total energy expenditure divided by basal metabolic rate (physical activity level, PAL) or as total energy expenditure minus basal metabolic rate (activity energy expenditure, AEE).

In free-living subjects the best method of assessing PAL and AEE is to combine measurements of total energy expenditure assessed using the doubly labeled water method with measurements of basal metabolic rate assessed using indirect calorimetry [3-5]. In the doubly labeled water technique [6], carbon dioxide production is estimated as the difference between the turnover rates of two tracer isotopes (^{2}H and ^{18}O) in the body water pool from which energy expenditure is calculated[6]. The method has been successfully validated in human subjects [6] and is unique since it can assess energy expenditure over a period of days and weeks in free-living individuals with minimal interference with daily life activities. However, this method of assessing PAL and AEE using the combination of the doubly labeled water method and indirect calorimetry was developed for clinical settings and is less suitable for larger population-based studies [4].

Other available methods of assessing energy expenditure due to physical activity are activity records, heart rate monitors, and accelerometers [4,5]. The accuracy of these methods is debated, and they are demanding for study participants and study personnel. The method most commonly used by epidemiologists is to administer paper questionnaires where participants report their physical activity retrospectively. Such questionnaires are easy to use and non-invasive [7-13] but are prone to bias since they rely on memory [14]. Neilson et al [15] reviewed a large number of studies where AEE estimates obtained using questionnaires had been compared with reference estimates based on the doubly labeled water method and indirect calorimetry. They concluded that most questionnaires are not sufficiently valid to assess AEE, indicating the need for new approaches within this area [15]. Telecommunication technologies have created possibilities for accurate real-time data collection and may thus reduce information bias due to retrospective data collection [16]. The high access to cell phones makes these devices potential tools for large-scale data collection in population-based studies, and previous studies have demonstrated the feasibility of collecting data using short message service (SMS), or text, messaging through cell phones [17,18]. SMS has been used in intervention trials, therapy management systems [19,20], as well as studies involving physical activity and weight loss [21-26]. With cell phones connected to the Internet, Web-like applications and more advanced questionnaires resembling Web-based questionnaires are also made possible [27-29].

The primary aim of this study was to validate PAL values aggregated over 14 days against reference estimates based on the doubly labeled water method and indirect calorimetry. The aim was also to evaluate the feasibility of collecting repeated measures of PAL through a Java-based questionnaire downloaded onto cell phones. Two traditional paper questionnaires were also used to compare the validity of the PAL estimates based on questionnaires administered daily over cell phones with the validity of the PAL estimates based on paper questionnaires administered retrospectively.

Methods

Subjects and Study Overview

Twenty-two healthy women were recruited to a study on energy metabolism and physical activity by means of advertisements in the local press between September 2007 and February 2008. The women were employed in different areas of work, such as office work, nursing, and childcare. On the morning of the first day, basal metabolic rate was measured. Each woman was then given a dose of stable isotopes and asked to collect urine samples during the subsequent 14-day period to measure total energy expenditure. (Measurement of metabolic rate, isotope administration, urine collection, and urine analysis are explained in more detail below.) A reference PAL (PAL_{ref}) covering this 14-day period was calculated as total energy expenditure divided by basal metabolic rate.

Each woman was instructed over the phone in how to download the Java-based questionnaire. Before the installation, the women were registered on a study-specific website. If a woman's cell phone didn't support the Java script, the study center lent her a phone that did. During the 14-day period, each woman received a text message every day at 9pm to remind her to answer the two questions about her physical activity during the same day through the Java-based questionnaire on her cell phone. The two questions were: (1) How physically active have you been during work/daytime today? and (2) How physically active have you been during leisure time/evening today? (Table 1). Answers to the questions were delivered to the study center in real-time through the cell phone's online connection and could be monitored on the study-specific website. If a woman did not reply, the study center sent her an additional SMS as a reminder. At the end of the 14-day period, each woman delivered the urine samples to the study center and while there answered two paper questionnaires regarding her physical activity during the preceding 14-day period. The study was designed according to the Helsinki Declaration and was approved by the Central Ethical Review Board in Stockholm, Sweden.

Table 1. The two questions administered using participants' cell phones

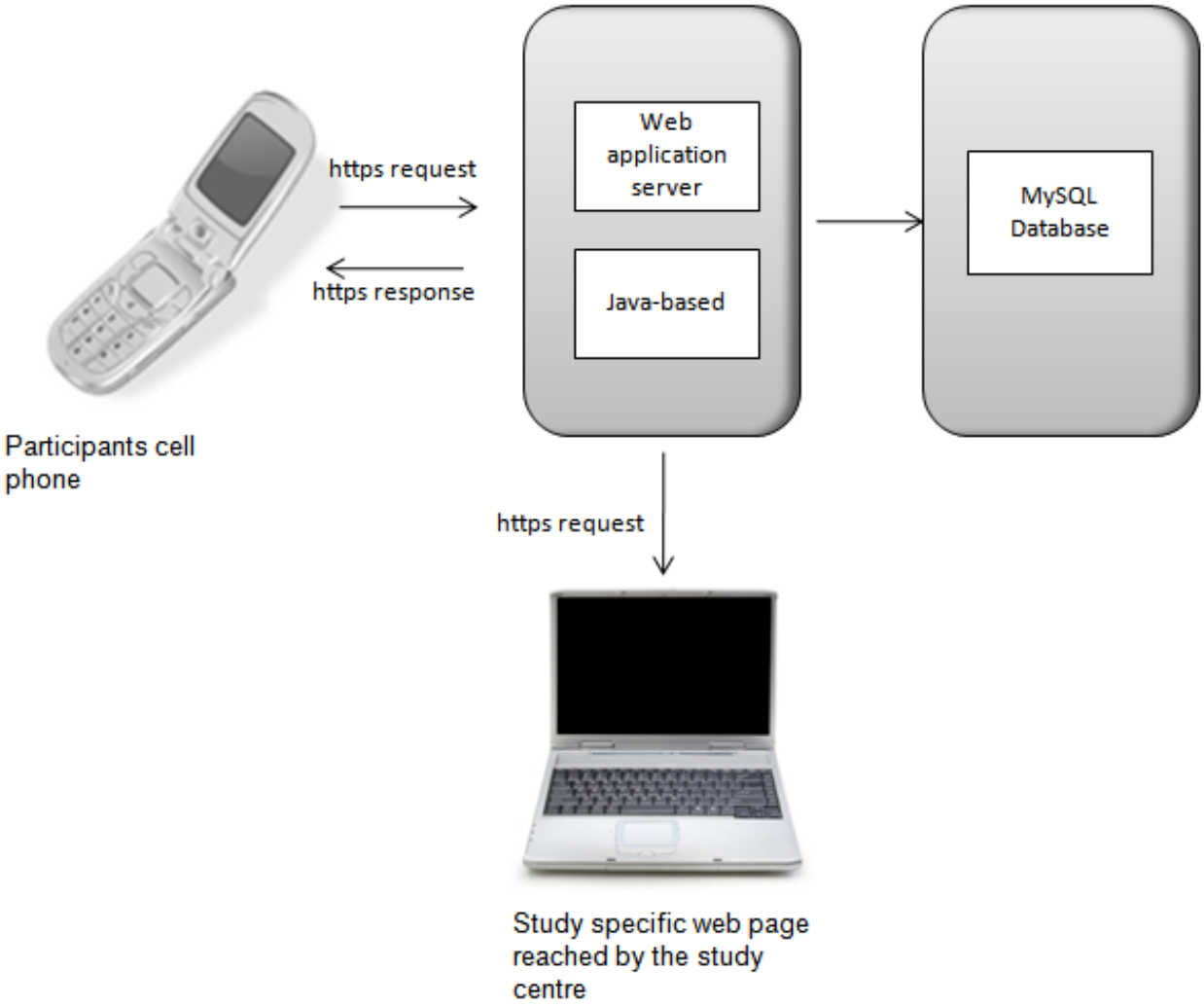
Answer Category	PAL Score
1. How physically active have you been during work/daytime today?	PAL: Derived Score for each Category[30][a]
a. Mostly sitting	1.55
b. Sitting/standing/walking	1.65
c. Standing/walking most of the time	1.85
d. Heavy work	2.2
2. How physically active have you been during leisure time/evening today?	Additional Contribution to PAL [31][a]
a.Mostly sitting	+0
b. Light/walking 30min	+0.06
c. Moderate/cycling≥30min	+0.15
d. Sport/cycling≥60min	+0.29

[a]See text for an explanation of how PAL was calculated from cell phone questions

Cell Phone Application

The download of the Java-based cell phone application was initialized through a WAP-push message. The Web client is built on Ajax in the Wicket application. The Web and application server was developed on Spring to build and run Java application and Hibernate to perform object relation mapping and query to the database. The database for collected information was MySQL, and Linux was used as the operating system. All data traffic was through https (hypertext transfer protocol secure). The procedure is described in Figure 1.

Figure 1. Schematic figure of the data communication between the cell phone, web application server, study-specific webpage, and the database



Measurement of basal metabolic rate and total energy expenditure

Basal metabolic rate was measured in the following manner. On the morning of the first day of the study, each woman's carbon dioxide production and oxygen consumption were measured after an overnight fast and 45 minutes of rest by indirect calorimetry during a 20-minute period using a ventilated hood system (Deltratrac Metabolic Monitor, Datex Instrumentarium Corp, Helsinki, Finland). The women came to the hospital by car to keep physical activity to a minimum. Carbon dioxide production and oxygen consumption were converted to basal metabolic rate through the equation by Weir [32].

To obtain measures of total energy expenditure, each woman was given an accurately weighed dose of the stable isotopes deuterium and oxygen-18 (0.09 g^2H_20 per kg body weight and 0.23 g $H_2{}^{18}0$ per kg body weight) after having collected two to three background urine samples. Women collected another six urine samples on days 1, 4, 6, 8, 11, and 15 after the day of dosing. Isotopic enrichments of dose and urine samples were analyzed as described by Lof et al [33]. Analytical precision for results expressed in ppm was 0.22 for ^2H and 0.03 for ^{18}O. Carbon dioxide production was calculated as described by Coward [34] with deuterium and oxygen-18 dilution spacesas the distribution spaces matching their respective rate-constants for isotopic disappearance and assuming 30% of water losses to be fractionated. The ratio between deuterium-space and oxygen-18 space was 1.033 ± 0.006. Total metabolic rate was calculated from carbon dioxide production assuming the food quotient to be 0.85 [35]. For one subject, the same dose and urine samples were analyzed nine times. The following coefficients of variation were obtained: 1.2% for total energy expenditure, 0.3% for total body water, and 0.15% for the ratio between deuterium-space and oxygen-18 space.

Calculation of PAL From Cell Phones

Women's answers to the two short questions about physical activity during the same day were converted to PAL according to the values shown in Table 1 that corresponded to participants' responses. All data were displayed on the study-specific website and stored in the local database. The two questions were derived in order to cover work/daytime and leisure time/evening activities. To our knowledge, these specific questions have not been used before, but a similar approach has been used in other studies [36,37]. PAL was calculated daily by adding the PAL value obtained from question 1 to the contribution of PAL from question 2. The PAL values for each option during work/day time (question 1) were obtained from Black et al [30], and the contribution to PAL from activities during leisure/evening time (question 2) was calculated from published values for energy costs for walking and cycling [31].Black et al derived the PAL values for different categories of work from 500 measurements based on the doubly labeled water method and indirect calorimetry [30]. The published energy costs (so called metabolic energy transfer, or MET, values) are total energy expenditure when walking and cycling divided by basal metabolic rate. These values were derived from measurements of energy expenditure using indirect calorimetry in laboratory settings [31]. PAL_{cell} was calculated as the mean of the 14 PAL values for days 1 through 14.

Calculation of PAL From the Two Paper Questionnaires

Questionnaires 1 and 2 are displayed in Multimedia Appendix 1. Questionnaire 1 consisted of one simple question that has been used in previous epidemiological studies [38,39]. Each woman graded her physical activity level during the last two weeks between 1 and 10, where 1 is very low, and 10 is very high. Women were informed that the value 1 should be interpreted as a sedentary lifestyle, while the value 5 represented a few long walks per week, and the value 10 represented exercise several times a week. The answer was converted to PAL_{quest1} . Value 1 represented a PAL value of 1.3. Each step up to 10 represented a 0.1 increase up to 2.2.

Questionnaire 2 required each woman to report the number of hours and minutes spent in nine different activity categories with assigned MET values [38] during one average 24-hour period during the last two weeks [8]. PAL_{quest2} was calculated as MET times the number of minutes for each activity category divided by 1440 minutes.

Calculation of AEE

For each woman, PAL_{cell}, PAL_{quest1}, and PAL_{quest2} were converted to corresponding measures of activity energy expenditure (AEE), that is AEE_{cell}, AEE_{quest1}, and AEE_{quest2}, and were calculated as (PAL times basal metabolic rate) minus basal metabolic rate. Reference estimates of AEE (AEE_{ref}) were calculated as total energy expenditure minus basal metabolic rate.

Statistical Analysis

Values are given as mean ± SD. Significant differences between mean values were identified by repeated measures analysis of variance (ANOVA) with subsequent post hoc analyses using Tukey's multiple comparison tests after having ascertained that the values were normally distributed. The agreement between PAL_{cell}, PAL_{quest1} , and PAL_{quest2} versus PAL_{ref} was evaluated using the Bland and Altman procedure [40]. This procedure is used to evaluate the agreement between estimates obtained using an alternative method and estimates obtained using a reference method. According to the procedure, the difference between the estimates obtained from the alternative and reference methods (y) was plotted against the average of these two estimates (x) for all subjects. The mean difference ± 2SD for all subjects was then calculated. Two SD was chosen as the measure of variation since this estimate corresponds to the limits of agreement used by Bland and Altman. To test for trend within methods, a linear regression model was fitted between x and y. The same procedure was used for AEE values. Intraclass correlation was used to measure intraindividual variation of daily physical activity levels through cell phones. All statistical tests were done on a two-sided 5% level of significance. All statistical analyses were performed using SAS 9.1.3 (SAS Institute Inc, Cary, NC, USA).

Results

Characteristics of Participants

Characteristics of women who participated in the study are shown in Table 2. The women represented a wide range of age, body mass index (BMI), and body weight. Of the 22 participating women, 4 had a BMI greater than 25 kg/m^2, and 2 had a BMI greater than 30 kg/m^2. None of the women were smokers, and all but 4 had a university or college degree. At baseline, 5 women (23%) reported that they never exercised, 5 women (23%) reported that they exercised 1 to 2 times per week, while 12 women (54%) reported that they exercised 3 times a week or more.

Table 2. Characteristics of the 22 women in the study

Characteristic	Mean ± SD	Range	Median	Q1[a]	Q3[b]
Age (years)	35.1 ± 8.3	20-45	37	29	42
BMI (kg/m^2)	23.7 ± 3.8	17.7-33.6	22.1	20.5	25.1
Height (m)	1.69 ± 0.06	1.55-1.81	1.69	1.65	1.74
Bodyweight (kg)	67.2 ± 13.3	47-102	65.5	58.7	73.4
Exercise (hours/week during the last year)[c]	2.4 ± 1.8	0-5	3	1	4

[a]First quartile: cutoff for lowest 25%

[b]Third quartile: cutoff for lowest 75%

[c]Defined as regular exercise (eg, sports, aerobics or running) during the year before the study. Walking or cycling for transportation (eg, to or from work) was not included.

Procedure

Of the 22 women, 14 (64%) had a cell phone that supported the Java-based application, and 8 (36%) borrowed a cell phone from the study center. During the 14-day period, all women but 2 answered all questions. Of the 2 women who did not answer all questions, one failed to report once, and the other, twice. For these women, $PAL_{cell}$was based on 13 and 12 days of reporting, respectively. None of the women asked to terminate the reporting before the end of the 14-day period, and at the end of the study, all women expressed little or no burden of completing the reporting during the 14-day period.

PAL Assessed Using Cell Phones

Table 3 demonstrates total energy expenditure and basal metabolic rate as well as PAL obtained using different methods. On average, PAL_{cell} was 1.82. This value was not statistically significantly different from PAL_{ref}, which was 1.83. The Bland and Altman plot for PAL_{cell} in comparison with PAL_{ref} is shown in Figure 2. The mean difference for PAL_{cell} and PAL_{ref} was small (0.014), and the limits of agreement were 2SD = 0.29. The test for trend was not statistically significant. The regression equation was y = -0.58 x + 1.05;r = -0.38;P = .08.

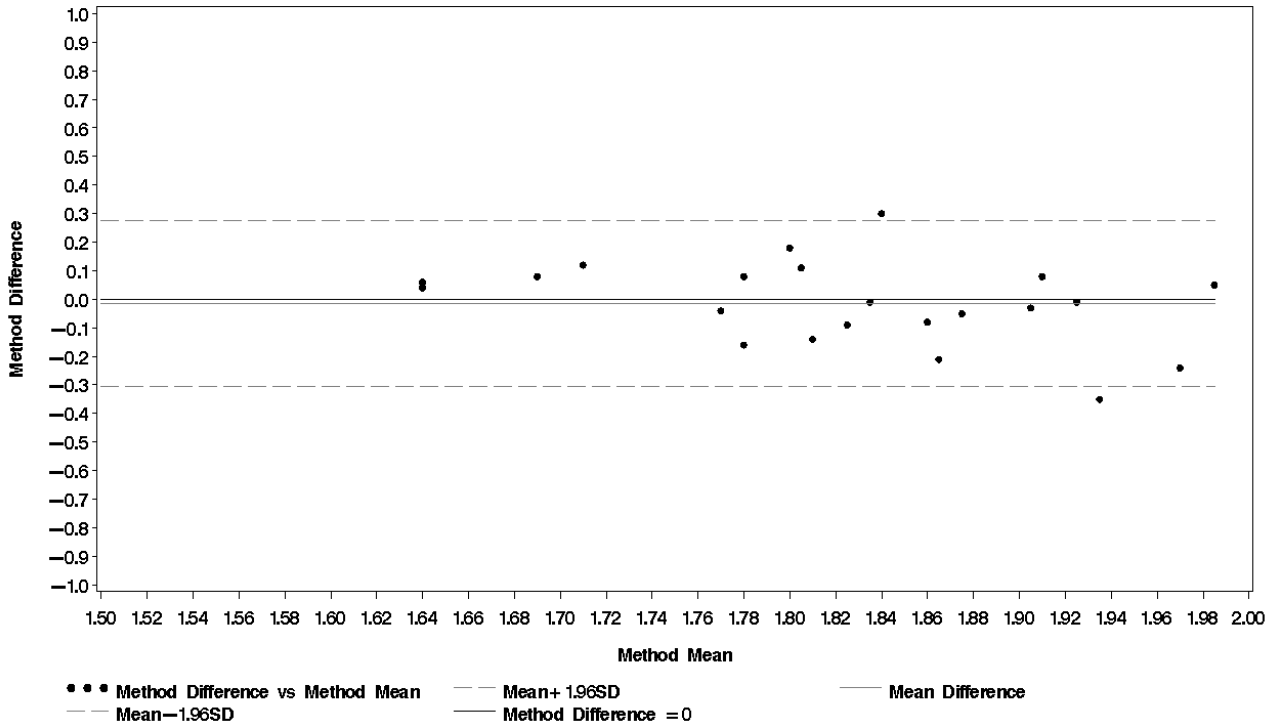
Table 3. Total energy expenditure,basal metabolic rate, and physical activity level (PAL) obtained using different methods in 22 Swedish women

Measurement	Mean ± SD	Range	Median	Q1[a]	Q3[b]
Total energy expenditure(kJ/24h)	10810 ± 1410	8130-13120	10640	9960	11620
Basal metabolic rate(kJ/24h)	5900 ± 710	4920-7950	5840	5490	6250
PAL_{ref}	1.83 ± 0.14	1.61-2.11	1.86	1.71	1.92
PAL_{cell}	1.82 ± 0.10	1.66-2.01	1.82	1.75	1.89
PAL_{quest1}	1.84 ± 0.23	1.50-2.20	1.85	1.60	2.00
PAL_{quest2}	1.90 ± 0.43	1.47-3.01	1.75	1.54	2.18

[a]First quartile: cutoff for lowest 25%

[b]Third quartile: cutoff for lowest 75%

Figure 2. Bland and Altman plot comparison of physical activity level obtained using cell phones during 14 days (PAL_{cell}) and physical activity level obtained using a combination of the doubly labeled water method and indirect calorimetry (PAL_{ref})



PAL Assessed Using Paper Questionnaires

The average $PAL_{quest1}$was 1.84 while the average $PAL_{quest2}$was 1.90 (Table 3). None of these values was statistically significantly different from PAL_{ref}. The Bland and Altman plots for $PAL_{quest1}$and PAL_{quest2}in comparison with $PAL_{ref}$are shown in Figure 3 and Figure 4. The mean difference was small for both $PAL_{quest1}$and PAL_{quest2}compared with PAL_{ref}(0.004 and 0.07 respectively). However, the limits of agreement were wider than for PAL_{cell} (PAL_{quest1} , 2SD = 0.51 and PAL_{quest2} , 2SD = 0.90). The test for trend was statistically significant for PAL_{quest1}(the regression equation was y = 0.79 x - 1.45; r = 0.44; *P*= .04) as well as for PAL_{quest2}(the regression equation was y 1.58 x - 2.88; *r*= 0.65; *P*< .001). Thus, both questionnaires (in particular questionnaire 2) overestimated higher PAL values while they underestimated lower PAL values.

Figure 3. Bland and Altman plot comparison of physical activity level obtained using the first questionnaire (PAL_{quest1}) and physical activity level obtained using a combination of the doubly labeled water method and indirect calorimetry (PAL_{ref})

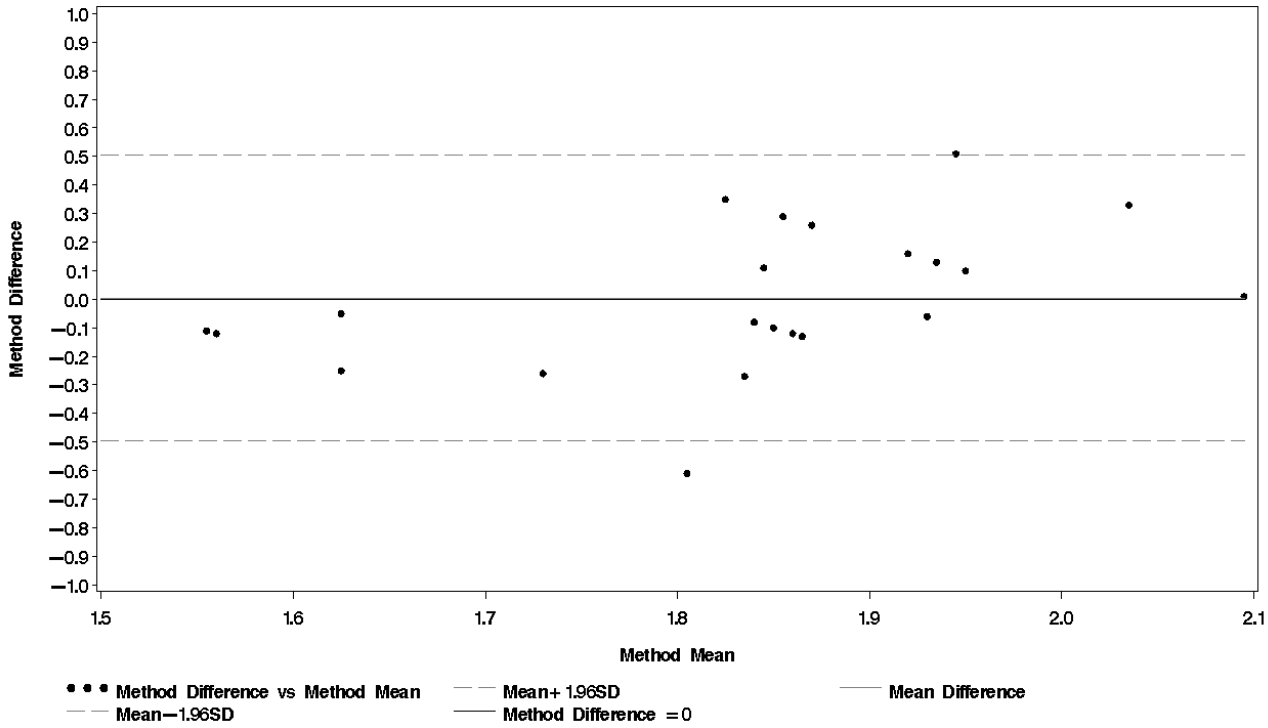
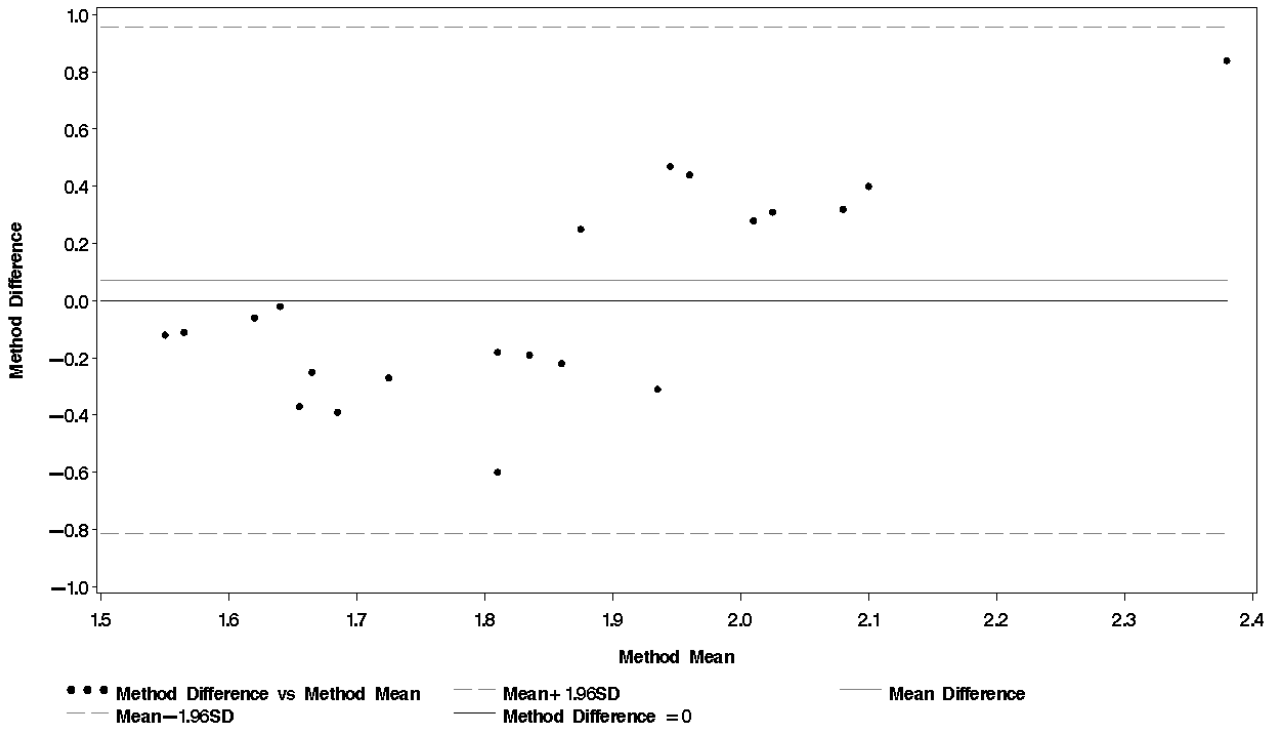


Figure 4. Bland and Altman plot comparison of physical activity level obtained using the second questionnaire (PAL_{quest2}) and physical activity level obtained using a combination of the doubly labeled water method and indirect calorimetry (PAL_{ref})



AEE From Cell Phones and Paper Questionnaires

AEE_{cell}, AEE_{quest1}, and AEE_{quest2}versus AEE_{ref} are shown in Table 4. The results were very similar to results for the PAL estimates. Similar results were obtained when the results were expressed as AEE in kJ/24h and per kg body weight (data not shown).

Table 4. Comparison of activity energy expenditure (AEE) assessed by cell phones and questionnaires in relation to reference estimates in 22 Swedish women

Measurement	Mean Difference (kJ/24h)	2SD for the Mean Difference (kJ/24h)	*r* [e]	*P* [f]
AEE_{cell} [a]- AEE_{ref} [b]	95	2380	-0.11	0.64
AEE_{quest1} [c]- AEE_{ref}	200	3630	0.42	0.05
AEE_{quest2} [d] -AEE_{ref}	540	4980	0.73	< .001

[a]AEE_{cell}= Activity energy expenditure obtained using the cell phone

[b]AEE_{ref}= Total energy expenditure minus basal metabolic rate

[c]AEE_{quest1}= Activity energy expenditure obtained using questionnaire 1

[d]AEE_{quest2}= Activity energy expenditure obtained using questionnaire 2

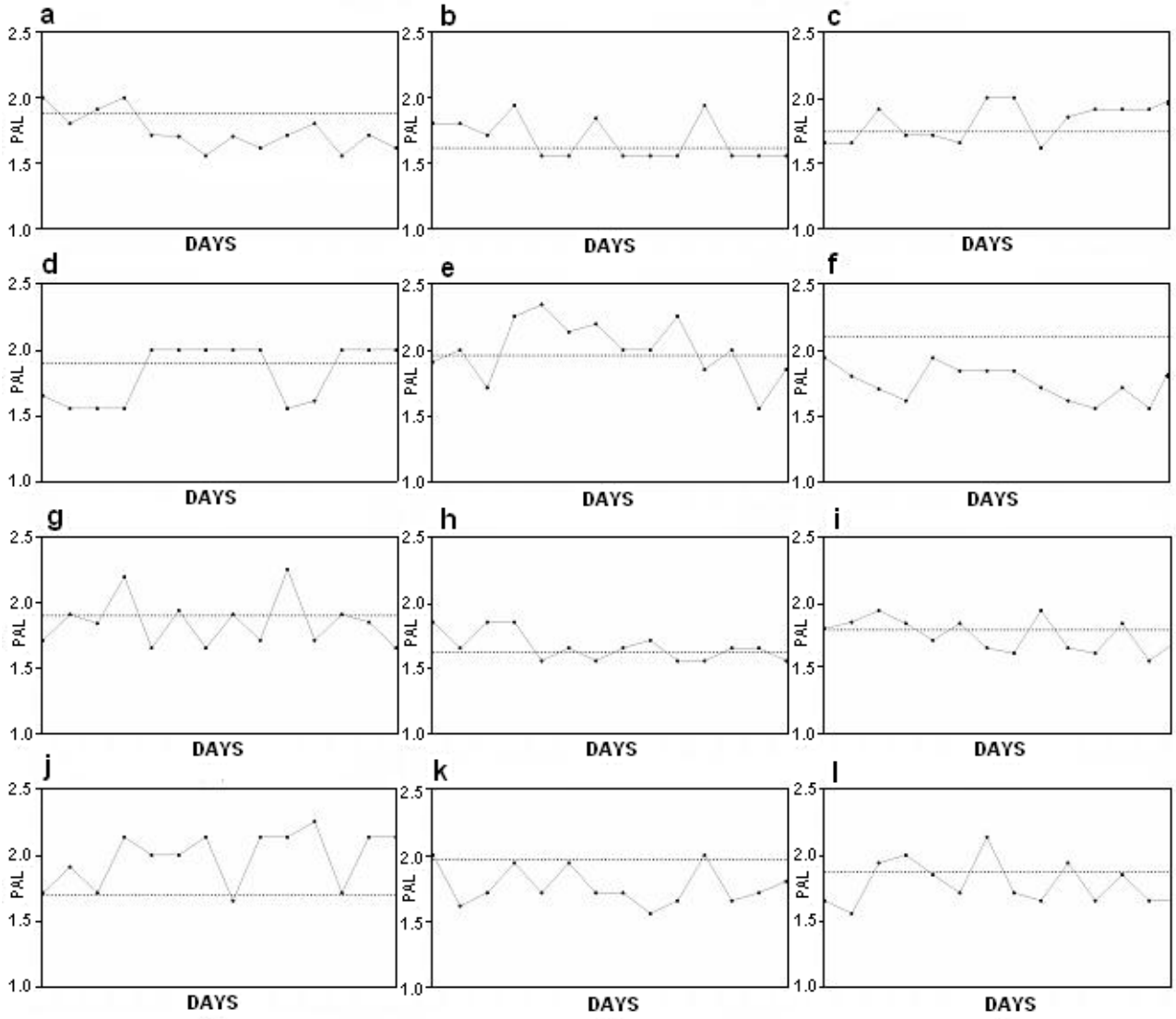
[e]The correlation coefficient for the linear regression between the average of the alternative method and the reference method and the difference between them. For instance, the *r* value when AEE_{cell}- AEE_{ref} was regressed on the average of AEE_{cell} and AEE_{ref}.

[f]*P*-value for the *r*-value

Day-to-Day Variation in PAL Obtained Using Cell Phones

PAL obtained using cell phones varied considerably from day to day during the 14-day study period (Figure 5). The intraclass correlation coefficient for the 22 women was estimated to be 0.20; thus about 20% of the variation is between women, while about 80% of the variation is due to day-to-day differences.

Figure 5. Daily PAL values obtained using cell phones during day 1 through day 14 for twelve selected women compared with PAL_{ref} (covering the whole 14-day period and shown as a straight dotted line for each woman)



Discussion

This study describes a novel approach to collecting data on physical activity using a Java-based physical activity questionnaire administered repeatedly through cell phones. The results indicate that measuring physical activity through cell phones is a promising method of assessing PAL that could be used in large-scale epidemiological studies.

The method generated high compliance and high acceptance among the participants. On average, PAL obtained using cell phones agreed well with reference estimates of PAL obtained using the doubly labeled water method and indirect calorimetry. Also the PAL values assessed by means of the two paper questionnaires were in good agreement with reference estimates. However, the limits of agreement for the difference between PAL obtained by cell phone and reference PAL were narrow (2SD = 0.29), while the corresponding limits for the two paper questionnaires were much wider (2SD = 0.51 for questionnaire 1 and 2SD = 0.90 for questionnaire 2). Furthermore, both questionnaires produced biased results (especially questionnaire 2), overestimating PAL of physically active women while underestimating PAL of less active women.

Only two previous studies have compared PAL estimates obtained from paper questionnaires to reference estimates for healthy adults based on the doubly labeled water method and indirect calorimetry. In these studies, paper questionnaires underestimated PAL by 0.12 units (6%) [33] or overestimated PAL by 0.31 units (31%) [41]. In our study, the cell phone estimates agreed by 0.01 units or 1% compared with the reference estimates on average. Only one of the former studies reported 2SD of the difference between PAL obtained using paper questionnaires and reference estimates. Their limits of agreement were wider than for the cell phone estimates in this study (2SD being 0.64 compared with 0.29) [33].

When expressing the results as AEE, the cell phone questions overestimated reference estimates of AEE by only 2% on average. For comparison, in the recent review by Neilson et al [15], only eight of twenty studies reported a mean difference in total energy expenditure or AEE less than 10%, and only four reported a mean difference less than 2% compared with reference estimates. Two SD was 2380 kJ/24h for the difference

between AEE_{cell} compared with AEE_{ref}. These limits of agreement are narrower than for most paper questionnaires that have been evaluated previously using reference estimates [15].

Our results showed a minor tendency (although not statistically significant) for PAL to be underestimated for physically active women using the cell phone questionnaire (Figure 2). This may be explained by the fact that in the current version, leisure activities included only walking, cycling, and sports. Other common everyday activities like gardening, moving, or playing with children may also increase PAL. Such activities could be included in a future version to test if they improve the ability of the cell phone questions to assess PAL of individual women. But the underestimation might also be explained by a skewed scale in the cell phone questionnaire, and that the options that require a higher level of physical activity should be given higher values.

Many paper questionnaires are designed to assess AEE. When comparing AEE between individuals, it should be expressed in kilojoules per kg body weight since smaller individuals have smaller AEE. This is a concern in epidemiological studies where body weight is often self-reported, creating a risk of bias. PAL is independent of body weight. In this study, PAL was chosen since our overall goal was to develop a procedure for use in epidemiological settings. Reference estimates of PAL were assessed as total energy expenditure divided by basal metabolic rate. Total energy expenditure consists of basal metabolic rate (50-70%), energy expenditure due to physical activity (20-40%), and diet-induced thermogenesis (5-10%). Thus, in the calculations of PAL, diet-induced thermogenesis was included. However, the cell phone and questionnaire estimates of PAL were calculated using published values for PAL [30] and MET [31], which also included diet-induced thermogenesis.

Figures 3 and 4 indicate that both questionnaires overestimated PAL of active women and underestimated PAL of more sedentary women. These results were obtained using the average of the alternative and reference method on the x-axis as described by Bland and Altman [40]. However, in some studies only the standard method is used on the x-axis. Had this procedure been used, the correlations would be somewhat different. According to Bland and Altman, using only the reference value will likely show a correlation whether there is a true association between difference and magnitude or not [42]. Thus, in this study, the average value was used on the x-axis.

Only 2 of the 22 women missed a reporting occasion (one failed to report once, the other, twice), indicating a high compliance for the study method. The high compliance may be due to the SMS sent as a reminder. The technique also provides a possibility for the study center to monitor whether the participant answers the cell phone questions during the study period. These features are an advantage compared with ordinary paper-based physical activity records. None of the 22 women expressed any concerns or that having to report daily had interfered with daily life.

The two questions were downloaded onto cell phones using a Java-based application; thus the participants had to have access to a cell phone with Java support. Out of the 22 women, 14 (64%) were able to download the Java-based questionnaire onto their own cell phone while 8 (36%) borrowed a cell phone from the study center. The same distribution of access to cell phones supporting a Java-based questionnaire was seen in another unpublished study. The advantage with Java in Web-based applications is that more advanced applications can be used, and the cost of data transfer is lower than the cost of sending an SMS. Furthermore, the Java program can also store data if the online connection is poor and then send the data once the signal is stronger. Once the program had been downloaded on the participant's cell phones, the time and effort from the study center was minimal as all communications were automatically managed by the application server. As the cell phone questionnaire included only two questions, the questions could be asked through SMS or interactive voice response (IVR) in order to increase the accessibility, though these techniques allow less advanced interface and are more dependent on connection to the cell network. However, both these techniques are accessible to all cell phones and have been used to communicate with patients in intervention and therapy studies [17,21,43].

The strengths of this study are that this is the first study to have developed and evaluated a cell phone based method of assessing PAL, and that gold standard methods were used as reference methods. One possible limitation of the PAL cell phone estimates is that they were based on self-report. Still, the mean difference ± 2SD between PAL obtained using cell phones and reference estimates was smaller than earlier studies that reported evaluations of PAL or total energy expenditure that had been assessed using objective methods like accelerometry [33,44]. The present study is small, with only 22 participants, but the study participants represented a wide range of BMI, exercise habits, and age. They were also a group of well-educated, highly motivated, and moderately active women. The results might not be applicable to other populations or subgroups in other social settings with a different age distribution or different levels of physical activity.

In this study high within-subject variation in relation to between-subject variation was noted for the cell phone estimates. The high within-subject variability is a natural consequence (although not shown earlier in this way) of subjects having different levels of physical activity from day to day, but it may also indicate that the cell phone questionnaire may not have captured all of the information about subjects' physical activity. We aimed to develop a simple procedure to classify physical activity levels of subjects suitable for epidemiological studies that could predict health outcomes many years later. For this purpose, short-term (day-to-day or hour-to-hour) variation is likely not a problem. However, our procedure may need modification to be useful for intervention studies since such studies may require more detailed information, such as the type and duration of physical activities performed. We cannot exclude the possibility that the low between-subject variation was to some extent due to the fact that our study participants were a relatively small group of women. However, this study is a first step toward developing and evaluating a unique cell phone strategy to assess PAL. These first results in a group of healthy Swedish women are promising, and the next phase should

include evaluations in other populations including men and different age groups.

Using the procedure described by Bland and Altman, the cell phone estimates produced narrow limits of agreement when compared with reference estimates. We can only speculate why, but it may be due to the fact that the cell phone questions were answered every evening, reducing the reporting error due to loss of memory. Another reason may be that the procedure takes advantage of a technique (cell phones) that is associated with "instant answers" which also may reduce reporting error. Paper questionnaires may provide subjects with more time to reflect on their answers and also to consider their total activity level, which may make subjects more prone to adjust their answers in order not to be "too inactive," for example.

Noteworthy is that the more detailed paper questionnaire used in this study was not superior to the use of only one rating question in assessing total physical activity. This finding supports the result from a recent study in rheumatoid arthritis patients, where good agreement was found for PAL values obtained through two questions on physical activity compared with reference estimates [36]. Future studies should evaluate if these results are also valid for other populations including men and other age groups.

In this study, large day-to-day variations in PAL were shown, indicating the need for repeated measurements within subjects to decrease this variation. When repeating the questions, it was possible to obtain average PAL estimates typical for individuals. Furthermore, this study shows that cell phones are useful tools for repeated collection of PAL values. This procedure for gathering information about physical activity has a great potential for large-scale prospective epidemiological studies. The cell phone-based procedure could also be used for data collection of other variables in epidemiological settings where variation could be decreased by repeated measurement, such as assessment of energy intake or other health-related variables.

In conclusion, a Java-based physical activity questionnaire administered using cell phones produced average PAL estimates that agreed well with PAL reference values. Furthermore, the limits of agreement between PAL obtained using cell phones and reference values were narrower than for the corresponding estimates obtained using paper questionnaires. Java-based questionnaires downloaded onto cell phones may be a feasible and cost-effective method of data collection for large-scale prospective studies of physical activity.

Acknowledgments

The study center would like to thank Novatelligence for Medipal and extend a special thank you to Johan Cederlund. The study was supported by Formas, the Swedish Research Council (15402, 4322), the Wallenberg Foundation, AFA Insurance, the Östergötland County Council, the M Bergvall's foundation, Thuring's foundation and the Swedish Society of Medicine.

Multimedia Appendix

The two paper questionnaires used in the study

[PDF file (Adobe PDF),82 KB - jmir_v12i1e2_app1.pdf]

References

1. Giovannucci E. Diet, body weight, and colorectal cancer: a summary of the epidemiologic evidence. J Womens Health (Larchmt) 2003 Mar;12(2):173-182. [Medline: 12737716] [doi: 10.1089/154099903321576574]
2. Mora S, Cook N, Buring JE, Ridker PM, Lee IM. Physical activity and reduced risk of cardiovascular events: potential mediating mechanisms. Circulation 2007 Nov 6;116(19):2110-2118 [FREE Full text] [Medline: 17967770] [doi: 10.1161/CIRCULATIONAHA.107.729939]
3. Schoeller DA, Ravussin E, Schutz Y, Acheson KJ, Baertschi P, Jequier E. Energy expenditure by doubly labeled water: validation in humans and proposed calculation. Am J Physiol 1986;250(5 Pt 2):R823-R830. [Medline: 3085521]
4. Ainslie P, Reilly T, Westerterp K. Estimating human energy expenditure: a review of techniques with particular reference to doubly labelled water. Sports Med 2003;33(9):683-698. [Medline: 12846591] [doi: 10.2165/00007256-200333090-00004]
5. Lagerros YT, Lagiou P. Assessment of physical activity and energy expenditure in epidemiological research of chronic diseases. Eur J Epidemiol 2007;22(6):353-362. [Medline: 17599242] [doi: 10.1007/s10654-007-9154-x]
6. ; International Atomic Energy Agency. Assessment of Body Composition and Total Energy Expenditure in Humans Using Stable Isotope Technique. IAEA Human Health Series No 3. Vienna, Austria; International Atomic Energy Agency 2009.
7. Arvidsson D, Slinde F, Hulthèn L. Physical activity questionnaire for adolescents validated against doubly labelled water. Eur J Clin Nutr 2005 Mar;59(3):376-383. [Medline: 15536471] [doi: 10.1038/sj.ejcn.1602084]
8. Lagerros YT, Mucci LA, Bellocco R, Nyrén O, Bälter O, Bälter KA. Validity and reliability of self-reported total energy expenditure using a novel instrument. Eur J Epidemiol 2006;21(3):227-236. [Medline: 16547838] [doi: 10.1007/s10654-006-0013-y]
9. Starling RD, Matthews DE, Ades PA, Poehlman ET. Assessment of physical activity in older individuals: a doubly labeled water study. J Appl Physiol 1999 Jun;86(6):2090-2096 [FREE Full text] [Medline: 10368377]

10. Mahabir S, Baer DJ, Giffen C, Clevidence BA, Campbell WS, Taylor PR, et al. Comparison of energy expenditure estimates from 4 physical activity questionnaires with doubly labeled water estimates in postmenopausal women. Am J Clin Nutr 2006 Jul;84(1):230-236. [Medline: 16825700]
11. Conway JM, Irwin ML, Ainsworth BE. Estimating energy expenditure from the Minnesota Leisure Time Physical Activity and Tecumseh Occupational Activity questionnaires - a doubly labeled water validation. J Clin Epidemiol 2002 Apr;55(4):392-399. [Medline: 11927208] [doi: 10.1016/S0895-4356(01)00497-8]
12. Bonnefoy M, Normand S, Pachiaudi C, Lacour JR, Laville M, Kostka T. Simultaneous validation of ten physical activity questionnaires in older men: a doubly labeled water study. J Am Geriatr Soc 2001 Jan;49(1):28-35. [Medline: 11207839] [doi: 10.1046/j.1532-5415.2001.49006.x]
13. Hertogh EM, Monninkhof EM, Schouten EG, Peeters PH, Schuit AJ. Validity of the Modified Baecke Questionnaire: comparison with energy expenditure according to the doubly labeled water method. Int J Behav Nutr Phys Act 2008;5(1):30 [FREE Full text] [Medline: 18505554] [doi: 10.1186/1479-5868-5-30]
14. Blair SN, Haskell WL. Objectively measured physical activity and mortality in older adults. JAMA 2006 Jul 12;296(2):216-218. [Medline: 16835428] [doi: 10.1001/jama.296.2.216]
15. Neilson HK, Robson PJ, Friedenreich CM, Csizmadi I. Estimating activity energy expenditure: how valid are physical activity questionnaires? Am J Clin Nutr 2008 Feb;87(2):279-291 [FREE Full text] [Medline: 18258615]
16. Ekman A, Litton JE. New times, new needs; e-epidemiology. Eur J Epidemiol 2007;22(5):285-292. [Medline: 17505896] [doi: 10.1007/s10654-007-9119-0]
17. Anhøj J, Møldrup C. Feasibility of collecting diary data from asthma patients through mobile phones and SMS (short message service): response rate analysis and focus group evaluation from a pilot study. J Med Internet Res 2004 Dec 2;6(4):e42 [FREE Full text] [Medline: 15631966] [doi: 10.2196/jmir.6.4.e42]
18. Kuntsche E, Robert B. Short message service (SMS) technology in alcohol research--a feasibility study. Alcohol Alcohol 2009;44(4):423-428. [Medline: 19482879] [doi: 10.1093/alcalc/agp033]
19. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [Medline: 19382860] [doi: 10.1089/tmj.2008.0099]
20. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [Medline: 19135907] [doi: 10.1016/j.amepre.2008.09.040]
21. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [Medline: 17478409] [doi: 10.2196/jmir.9.2.e7]
22. Hurling R, Fairley B, Dias M. Internet-based exercise intervention systems: Are more interactive designs better? Psychol Health 2006;21(6):757-772. [doi: 10.1080/14768320600603357]
23. Joo NS, Kim BT. Mobile phone short message service messaging for behaviour modification in a community-based weight control programme in Korea. J Telemed Telecare 2007;13(8):416-420. [Medline: 18078554] [doi: 10.1258/135763307783064331]
24. Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [Medline: 19141433] [doi: 10.2196/jmir.1100]
25. Newton KH, Wiltshire EJ, Elley CR. Pedometers and text messaging to increase physical activity: randomized controlled trial of adolescents with type 1 diabetes. Diabetes Care 2009 May;32(5):813-815 [FREE Full text] [Medline: 19228863] [doi: 10.2337/dc08-1974]
26. Gerber BS, Stolley MR, Thompson AL, Sharp LK, Fitzgibbon ML. Mobile phone text messaging to promote healthy behaviors and weight loss maintenance: a feasibility study. Health Informatics J 2009 Mar;15(1):17-25. [Medline: 19218309] [doi: 10.1177/1460458208099865]
27. Kollmann A, Riedl M, Kastner P, Schreier G, Ludvik B. Feasibility of a mobile phone-based data service for functional insulin treatment of type 1 diabetes mellitus patients. J Med Internet Res 2007;9(5):e36 [FREE Full text] [Medline: 18166525] [doi: 10.2196/jmir.9.5.e36]
28. Morak J, Schindler K, Goerzer E, Kastner P, Toplak H, Ludvik B, et al. A pilot study of mobile phone-based therapy for obese patients. J Telemed Telecare 2008;14(3):147-149. [Medline: 18430285]
29. 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] [Medline: 18550322] [doi: 10.1016/j.amepre.2008.05.001]
30. Black AE, Coward WA, Cole TJ, Prentice AM. Human energy expenditure in affluent societies: an analysis of 574 doubly-labelled water measurements. Eur J Clin Nutr 1996 Feb;50(2):72-92. [Medline: 8641250]
31. Ainsworth BE, Haskell WL, Whitt MC, Irwin ML, Swartz AM, Strath SJ, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000;32(9 Suppl):S498-S504. [Medline: 10993420]
32. Weir JB. New methods for calculating metabolic rate with special reference to protein metabolism. Nutrition 1949;6(3):213-221. [Medline: 2136000]
33. Lof M, Hannestad U, Forsum E. Comparison of commonly used procedures, including the doubly-labelled water technique, in the estimation of total energy expenditure of women with special reference to the significance of body fatness. Br J Nutr 2003 Nov;90(5):961-968. [Medline: 14667189] [doi: 10.1079/BJN2003975]

34. Coward WA. Stable isotopic methods for measuring energy expenditure. The doubly-labelled-water (2H2(18)O) method: principles and practice. Proc Nutr Soc 1988 Sep;47(3):209-218. [Medline: 3076004] [doi: 10.1079/PNS19880037]
35. Black AE, Prentice AM, Coward WA. Use of food quotients to predict respiratory quotients for the doubly-labelled water method of measuring energy expenditure. Hum Nutr Clin Nutr 1986 Sep;40(5):381-391. [Medline: 3771290]
36. Johansson G, Westerterp KR. Assessment of the physical activity level with two questions: validation with doubly labeled water. Int J Obes (Lond) 2008 Jun;32(6):1031-1033. [Medline: 18392036] [doi: 10.1038/ijo.2008.42]
37. Lof M, Forsum E. Validation of energy intake by dietary recall against different methods to assess energy expenditure. J Hum Nutr Diet 2004 Oct;17(5):471-480. [Medline: 15357701] [doi: 10.1111/j.1365-277X.2004.00554.x]
38. Trolle-Lagerros Y, Mucci LA, Kumle M, Braaten M, Weiderpass E, Hsieh CC, et al. Physical activity as a determinant of mortality in women. Epidemiology 2005;16(6):780-785. [Medline: 16222168]
39. Margolis KL, Mucci L, Braaten T, Kumle M, Trolle Lagerros Y, Adami HO, et al. Physical activity in different periods of life and the risk of breast cancer: the Norwegian-Swedish Women's Lifestyle and Health cohort study. Cancer Epidemiol Biomarkers Prev 2005 Jan;14(1):27-32 [FREE Full text] [Medline: 15668472]
40. 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]
41. Conway JM, Seale JL, Jacobs DR, Irwin ML, Ainsworth BE. Comparison of energy expenditure estimates from doubly labeled water, a physical activity questionnaire, and physical activity records. Am J Clin Nutr 2002 Mar;75(3):519-525 [FREE Full text] [Medline: 11864858]
42. 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] [doi: 10.1016/S0140-6736(95)91748-9]
43. Ostojic V, Cvoriscec B, Ostojic SB, Reznikoff D, Stipic-Markovic A, Tudjman Z. Improving asthma control through telemedicine: a study of short-message service. Telemed J E Health 2005 Feb;11(1):28-35. [Medline: 15785218] [doi: 10.1089/tmj.2005.11.28]
44. Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity (Silver Spring) 2007 Oct;15(10):2371-2379. [Medline: 17925461] [doi: 10.1038/oby.2007.281]

Abbreviations

AEE: activity energy expenditure
BMI: body mass index
https: hypertext transfer protocol secure
IVR: interactive voice response
Java: Java technology was created as a computer programming tool at Sun Microsystems in 1991
MET: metabolic energy turnover
PAL: physical activity level
SD: standard deviation
SMS: short message service

Edited by G Eysenbach; submitted 09.07.09; peer-reviewed by F Slinde, T Bickmore, E Hekler; comments to author 31.07.09; revised version received 21.10.09; accepted 01.12.09; published 29.01.10.

Please cite as:
Bexelius C, Löf M, Sandin S, Trolle Lagerros Y, Forsum E, Litton JE
Measures of Physical Activity Using Cell Phones: Validation Using Criterion Methods
J Med Internet Res 2010;12(1):e2
URL: http://www.jmir.org/2010/1/e2/
doi:10.2196/jmir.1298
PMID:20118036

Original Paper

How Breast Cancer Patients Want to Search for and Retrieve Information From Stories of Other Patients on the Internet: an Online Randomized Controlled Experiment

Regina Overberg[1], MSc; Wilma Otten[2], PhD; Andries de Man[1], PhD; Pieter Toussaint[3], PhD; Judith Westenbrink[4]; Bertie Zwetsloot-Schonk[1], MD, PhD

[1]Clinical Informatics Group, Leiden University Medical Center, Leiden, The Netherlands

[2]Medical Decision Making, Leiden University Medical Center, Leiden, The Netherlands

[3]Department of Computer and Information Science, Norwegian University of Science and Technology, Trondheim, Norway

[4]The Amazones Foundation, Haarlem, The Netherlands

Corresponding Author:
Regina Overberg, MSc
Clinical Informatics Group
Leiden University Medical Center
Postal Zone S5-P, PO Box 9600, 2300 RC Leiden
The Netherlands
Phone: +31 71 526 9700
Fax: +31 71 526 8280
Email: r.i.overberg@lumc.nl

Abstract

Background: Other patients' stories on the Internet can give patients information, support, reassurance, and practical advice.

Objectives: We examined which search facility for online stories resulted in patients' satisfaction and search success.

Methods: This study was a randomized controlled experiment with a 2x2 factorial design conducted online. We facilitated access to 170 stories of breast cancer patients in four ways based on two factors: (1) no versus yes search by story topic, and (2) no versus yes search by writer profile. Dutch speaking women with breast cancer were recruited. Women who gave informed consent were randomly assigned to one of four groups. After searching for stories, women were offered a questionnaire relating to satisfaction with the search facility, the stories retrieved, and impact of the stories on coping with breast cancer. Of 353 enrolled women, 182 (51.6%) completed the questionnaire: control group (n = 37), story topics group (n = 49), writer profile group (n = 51), and combination group (n = 45).

Results: Questionnaire completers were evenly distributed over the four groups ($\chi^2_3 = 3.7$, $P = .30$). Women who had access to the story topics search facility (yes vs no): were more positive about (mean scores 4.0 vs 3.6, $P = .001$) and more satisfied with the search facility (mean scores 7.3 vs 6.3, $P < .001$); were more positive about the number of search options (mean scores 2.3 vs 2.1, $P = .04$); were better enabled to find desired information (mean scores 3.3 vs 2.8, $P = .001$); were more likely to recommend the search facility to others or intend to use it themselves (mean scores 4.1 vs 3.5, $P < .001$); were more positive about how retrieved stories were displayed (mean scores 3.6 vs 3.2, $P = .001$); retrieved stories that better covered their information needs (mean scores 3.0 vs 2.6, $P = .02$); were more satisfied with the stories retrieved (mean scores 7.1 vs 6.4, $P = .002$); and were more likely to report an impact of the stories on coping with breast cancer (mean scores 3.2 vs 2.9, $P =. 02$). Three main effects were associated with use of the writer profile search (yes vs no): being more positive about (mean scores 3.9 vs 3.6, $P = .005$) and more satisfied with the search facility (mean scores 7.1 vs 6.5, $P =. 01$), and being more positive about how retrieved stories were displayed (mean scores 3.8 vs 2.9, $P < .001$). For satisfaction with the search facility, an interaction effect was found ($P = .03$): at least one of the two search facilities was needed for satisfaction.

Conclusions: Having access to the story topics search facility clearly had the most positive effect on patient satisfaction and search success.

(J Med Internet Res 2010;12(1):e7) doi:10.2196/jmir.1215

KEYWORDS

Breast cancer; life experiences; social support; Internet; information retrieval; patient satisfaction.

Introduction

Patients value having access to stories of other patients as it provides them with emotional support, information, reassurance, and practical advice [1]. The Internet is a valuable resource for accessing stories because of its privacy and 24-hour availability without the need to leave one's home [2]. Two well-known examples of web-based applications that include personal stories of patients are The Comprehensive Health Enhancement Support System (CHESS) [3] and The Database of Individual Patients' Experiences of Illness (DIPEx) [4].

Studies of online patient stories have focused on several factors. These include why patients publish their stories online and what this means in a broader sociological context [5-7]. Wise et al [8] found that accessing personal stories in a computer-based patient support system had a positive effect on patients' healthcare participation, which entailed participation preferences, confidence, and communication with their doctor. Little is known, however, about how patients search online for stories of other patients and whether they can find relevant ones.

Some qualitative studies have found that patients appreciate the ability to select stories of other patients of a particular age or who have opted for similar treatment [1,9]. In addition, searching by topics seems also to be of interest [10]. Some websites with patient stories provide a search facility to search for personal characteristics of the story writers and/or for topics written about in the stories [11-13]. However, to our knowledge, patients' satisfaction and search success with these search facilities have not yet been studied.

In the present study, we examined which search facilities for patient stories resulted in satisfaction with the search process and the stories retrieved. We also studied the impact of the stories retrieved on coping with illness. Our expectation was that having a search facility would be an improvement compared with not having a search facility. Moreover, we expected that a combination of search facilities would result in higher satisfaction than a single search facility because a combination of search facilities may result in more opportunities to find a relevant story.

Methods

Design and Procedure

Study Design

We focused our study on patients with breast cancer. We contacted the board of The Amazones Foundation, which was founded by a group of young women with breast cancer to provide their peers with information and support. The Amazones Foundation developed a website for young women with breast cancer [14] that provides information and advice, a calendar of activities, an online support group, and links to other sites. The website also has a section with personal stories. Women can anonymously submit their own story to the site. The stories are presented alphabetically by writers' nicknames. If a writer passes away, her story remains on the site accompanied by an obituary written by the website moderators.

We were granted permission by the board of The Amazones Foundation to conduct our study. In January 2007 we downloaded all 170 stories available at that time on their website for use in our study. We facilitated access to the stories in three ways: (1) with a search facility for story topics, (2) with a search facility for writer profiles, and (3) with a combination of these two search facilities. In addition, a control group could access the stories by means of the original alphabetical listing by story writer. We implemented these four ways of facilitating access to the stories on a separate study website. This resulted in four groups based on two independent factors: (1) no versus yes search by story topics, and (2) no versus yes search by writer profile (Table 1). In each of the four groups the same set of 170 stories could be searched. Figures 1 to 4 show screenshots of the search pages available to the four groups. We requested that the board of the Amazones Foundation not to participate in the study.

The present study is reported in accordance with the CHERRIES checklist, which is a checklist for reporting results of Internet e-surveys [15]. It was not registered as a clinical trial on ClinicalTrials.gov, a registry of clinical trials conducted around the world, because our study does not correspond to the definition of a clinical trial as provided in their glossary.

Table 1. The 2x2 factorial design of the study

		Writer Profile Search	
		No	Yes
Story Topics Search	No	Control group: see Figure 1	Writer profile group: see Figure 3
	Yes	Story topics group: see Figure 2	Combination group: see Figure 4

Recruitment Process

Recruitment announcements were disseminated online using banners on the websites of several Dutch patient and health organizations and offline using posters and flyers in waiting rooms of several hospitals. Dutch-speaking women with breast cancer were invited to participate irrespective of other personal characteristics. The offline recruitment announcements gave the URL of the study website, and the online announcements contained a hyperlink to the study website. The study was accessible to each visitor of the site, but only visitors who met the inclusion criteria were further directed to the informed consent page. After finishing the final questionnaire, participants

could send a ready-made email message with the URL of the study website to other women who might be interested in participation. The study website was accessible in the period June through November 2007. During this time period, women could choose for themselves on what day and time they wanted to participate. No incentives were offered for participation.

Study Website

The first page of the study website provided the following information about the study: study objective, information about the researchers, study inclusion criteria, details about participation, expected time within which to complete the study, and contact details. When a website visitor chose to participate by clicking the button "I would like to participate", two questions were presented to check whether the visitor met the inclusion criteria (ie, being female and having been diagnosed with breast cancer). If this was the case, the visitor was asked to read the informed consent statement. By agreeing with the informed consent statement, the visitor declared that her participation was voluntary, that she understood what participating entailed, and that she was aware of what data would be recorded. To agree with the informed consent statement, the visitor had to check the box "I agree" and then click on the button marked "Next". After this the visitor was asked whether she was certain that she agreed with the informed consent statement. In this way, we assured that women did not automatically agree to participate. Before the participant was randomly assigned to one of the four groups, she was asked to provide a short description of the information she wanted to search for in the stories. Random assignment of each participant to a study group was carried out by an algorithm that was part of the study website. The chance of being assigned to a group was equal for all four groups, that is, 1 in 4 (except when more than one session was conducted from the same IP address; see the section below labeled "Technical Aspects"). Once assigned to a group, a participant could search for and read the stories as long as she liked. When finished with searching and reading she was asked to complete a final questionnaire posted on the study website about satisfaction with the search process and the stories retrieved and the stories' impact on coping with breast cancer.

Figure 1. Screenshot of the control group search page

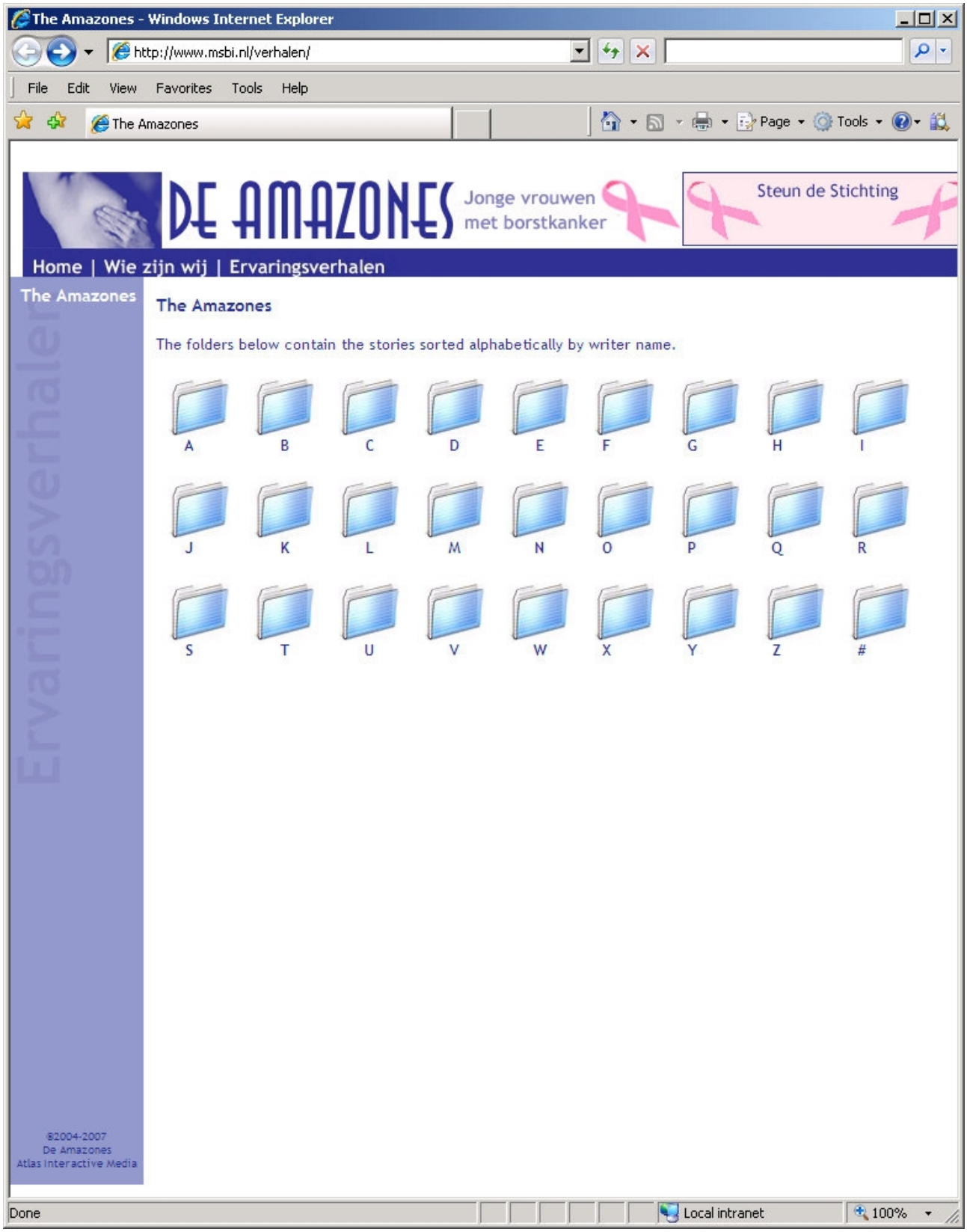


Figure 2. Screenshot of the story topics group search page

Figure 3. Screenshot of the writer profile group search page

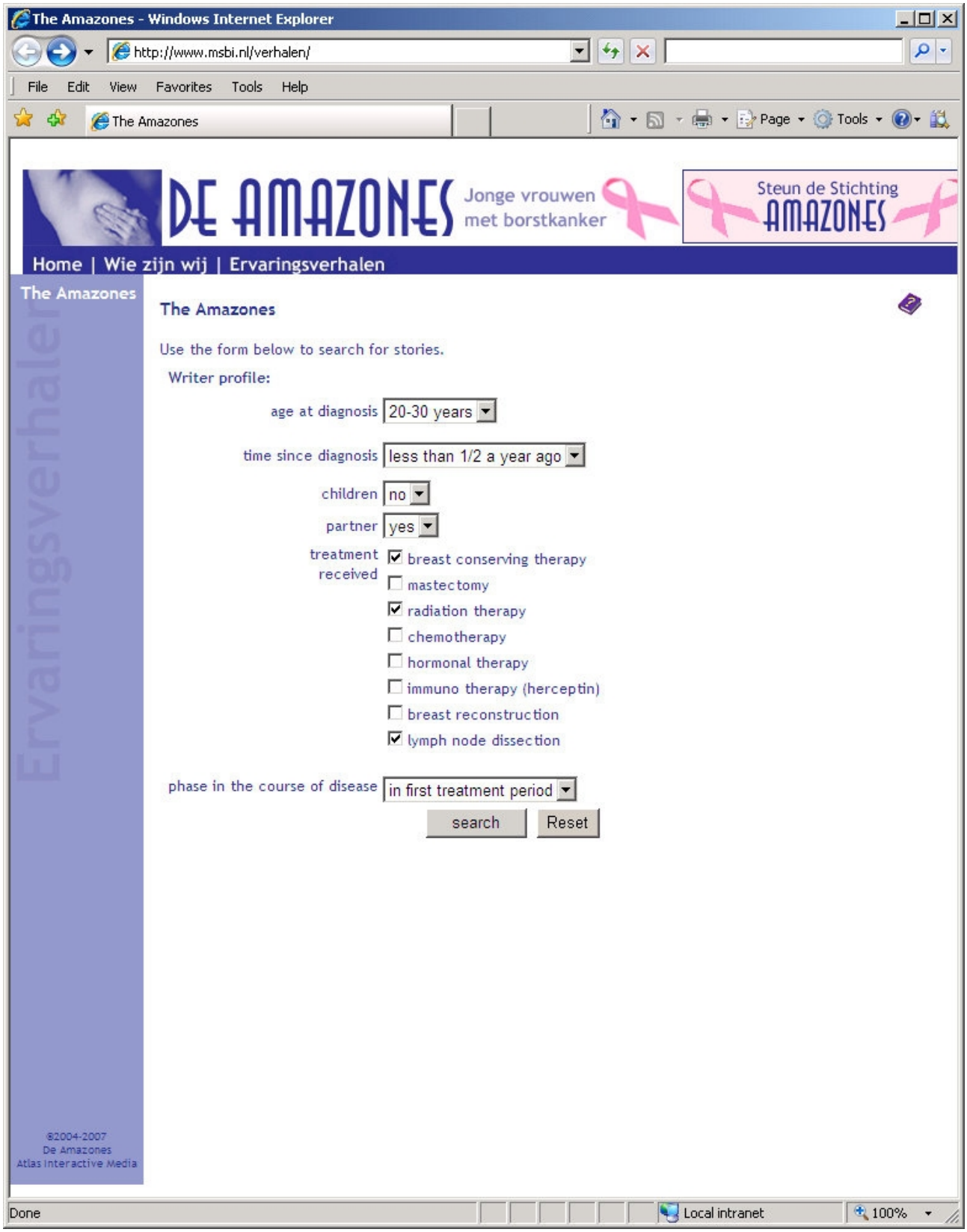
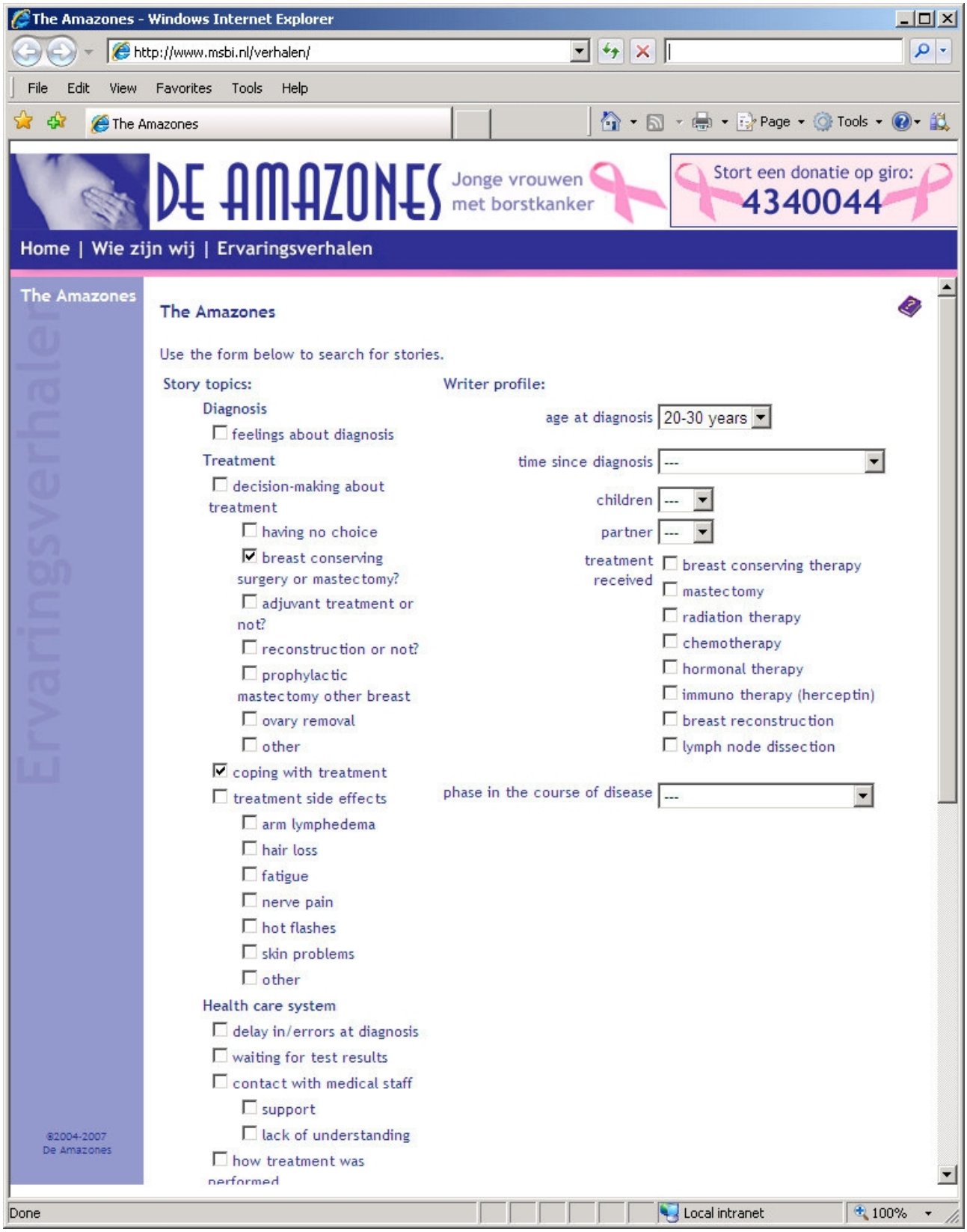


Figure 4. Screenshot of the combination group search page



Ethical Aspects

Participants remained anonymous since no log-in, name, or address were required. In order to minimize traces on each client's computer, no cookies were used. Recording of log data did not start until participants had agreed to the informed consent statement. Questionnaire responses were not saved until participants confirmed at the end of the final questionnaire that they agreed to submit their responses. Data were saved in a password protected SQL database only accessible to the researchers. Participants could stop at any time without receiving pop-ups or text when leaving the study website.

Before the start of the study, our research proposal (see Multimedia Appendix 1) was presented to the Ethical Committee of the Leiden University Medical Centre (archive number 06/43). The Committee concluded that our study involved no medical intervention and that we could proceed. Our intervention consisted of providing access to the stories already available on the website of the Amazones Foundation in several new ways.

Development of Intervention Groups

Search Facilities

To develop the search facilities, all 170 stories were coded according to a coding scheme for story topics and a coding scheme for writer profile (Table 2). The topics and personal characteristics in the coding schemes were chosen because they had been used on other websites that contained breast cancer stories [11-13] or by other authors of studies in this field [1,9,10]. For the characteristic "phase in the course of disease," the category "passed away" was assigned to stories that contained an obituary.

Participants could search for age using the categories: 20-30 years, 30-40 years, 40-50 years, and over 50 years. Participants could search for time since diagnosis using the categories: less than half a year ago, ½ - 1 year ago, 1-3 years ago, 3-5 years ago, and more than 5 years ago. To ensure that participants were aware of all search facilities, the search button was placed at the bottom of the page. In the groups with access to a single search facility, it was possible to search for more than one topic or more than one writer characteristic. In the combination group, participants could choose whether they wanted to search for story topics only, for writer characteristics only, or for both. Searching for more items simultaneously was based on the OR Boolean operator.

Table 2. Coding schemes for story topics and writer profile

Search Facility	Coding Scheme
Story topics (domains)	
Diagnosis	Feelings about diagnosis
Treatment	Decision-making about treatment: (1) having no choice, (2) breast conserving surgery or mastectomy?, (3) adjuvant treatment or not?, (4) reconstruction or not?, (5) prophylactic mastectomy of other breast, (6) ovary removal, (7) other
	Coping with treatment
	Treatment side effects:(1) arm lymphedema, (2) hair loss, (3) fatigue, (4) nerve pain, (5) hot flashes, (6) skin problems, (7) other
Health care system	Delay in/errors at diagnosis
	Waiting for test results
	Contact with medical staff: (1) support, (2) lack of understanding
	How treatment was performed: (1) satisfied, (2) dissatisfied
	Second opinion
Living with it	Work and insurances
	Family and friends: (1) support, (2) lack of understanding, (3) talking with and worrying about
	Body image and sexuality: (1) (partly) missing a breast, (2) partner's reaction
	Pregnancy issues: (1) pregnant at diagnosis, (2) wanting to become pregnant after treatments
	Coping with breast cancer: (1) thinking (emotional-focused coping), (2) doing (problem-focused coping)
	Practical advices
	Concerns about heredity
	Coping with metastasized breast cancer
Writer profile (personal characteristics)	
Age at diagnosis	Number of years
Time since diagnosis	Number of months
Partner	(1) No, (2) yes
Children	(1) No, (2) yes
Treatment received	(1) Breast conserving therapy, (2) mastectomy, (3) radiation therapy, (4) chemotherapy, (5) hormonal therapy, (6) immuno therapy (herceptin), (7) breast reconstruction, (8) lymph node dissection
Phase in the course of disease	(1) In first treatment period, (2) free of cancer, (3) cancer for second time, (4) metastasized cancer, (5) passed away

Weight Assignment in Story Retrieval

For every search performed by the participants, a weight between 0 and 1 was assigned to each of the 170 stories in the database. If a story matched exactly with the search objectives, it received a weight of 1. Story weights were calculated with every new search. Therefore, the weight assigned to a story could change with every search.

In the story topics group, weights were calculated by dividing the number of topics found in a story by the number of topics that were searched for. For example, when a participant searched for four topics, all stories containing one of these four topics received a weight of ¼ (0.25).

In the writer profile group, a weight was assigned to each of the personal characteristics that a participant searched for. These weights were then multiplied with each other to calculate the weight of a story as a whole.

If the age of a writer fell in the age category that the participant was searching for, then "age" received a weight equal to 1. The more the age of the writer deviated from the age category that the participant was searching for, the lower the weight that "age" received. In a similar way, weights for "time since diagnosis" were assigned.

If the partner status of a writer exactly matched the partner status that the participant was searching for, then "partner" received a weight equal to 1. If the partner status of a writer is unknown, then this characteristic received a weight of 0.5 irrespective of the partner status that the participant was searching for. If the partner status of a writer was the opposite of the partner status that a participant was searching for, then this characteristic received a weight of 0.2. We did not assign a weight of 0 for the latter case because then the weight for the whole story would be 0. In similar ways, weights for the other categorical variables were assigned.

In the combination group, weight assignment was similar to that of the previous two groups or a multiplication of these two, depending on whether a participant searched for story topics only, for writer profile only, or for both.

Number of Stories Retrieved

It was decided to present participants with at least ten stories after each search. The total number of stories presented after a search depended on the distribution of the weights assigned to the stories. All stories with the same weight as the tenth story were presented because we saw no valid reason for presenting only a portion of the stories with that weight. For example, when five stories matched exactly with the search objectives, that is, weight equal to 1, 20 stories received a weight of 0.80, 45 stories received a weight of 0.60, and 100 stories received a weight of 0.40, then 25 stories (5 + 20) would be presented. Accordingly, if no stories exactly matched the search objectives, still at least ten stories were presented. The list of retrieved stories showed the extent to which the stories matched the search objectives. In this way we tried to present participants with neither too few nor too many stories.When participants did not fill in the search page, no stories were presented to them because we wanted to ensure that participants were aware of the search facility.

Story Display and Sequence

The retrieved stories were displayed as a list giving for each story the writer's nickname and the story's weight. Weights were represented as a number of pink ribbons. In addition, the search criteria that were fulfilled were given in each group, that is, the topics found, the writer's characteristics, or both (see Multimedia Appendix 2). The list of stories was sorted by weight with the story with the heighest weight at the top. Stories with the same weight in the story topics group were displayed as follows. For each story, the percentage of text of the story relating to the topics the participant searched for was calculated. Stories with the highest percentage were ranked first. In the writer profile group, stories with the same weight were sorted by the age of the writers, and if age of writers was equal by time since diagnosis. Clicking on a story title from the list displayed the complete story.

Final Questionnaire

Demographic and Disease Characteristics

Participants were asked to provide demographic information such as age, marital status, children, religion, education, and employment status. They were also asked to report characteristics of their cancer, such as time since diagnosis, type of diagnosis, metastases in axillary lymph nodes or other parts of the body, treatment undergone, and prognosis.

Use of the Internet and the Amazones Website (Before Study Participation)

The participants were asked to indicate their frequency of Internet use, the type of activities in which they engage on the Internet, and whether they had read stories of other patients on the Internet before. Moreover, they were asked to report how often they had visited the Amazones website before, how familiar they were with this website, and how many of the stories on this website they read before.

Satisfaction With the Search Process, the Stories Retrieved, and the Stories' Impact on Coping With Breast Cancer

The constructs listed below were used to measure the three outcomes. Cronbach alphas were calculated using SPSS version 16.0 (SPSS Inc, Chicago, IL, USA) by conducting reliability analyses. Reverse phrased items were recoded. We found the internal consistency for each construct to be good or satisfactory (Cronbach alpha = 0.71 to 0.88). The items "overall satisfaction with the search facility" and "overall satisfaction with the stories retrieved" were answered using 10-point Likert scales; all other items were answered using 5-point Likert scales. For an overview of the items belonging to all the below mentioned constructs, see the Multimedia Appendix 3.

To measure satisfaction with the search process, 13 items were formulated (partially based on [16,17]). "Opinion about the search facility" was measured using 5 items (Cronbach alpha = .88). To measure "the extent to which the search options enabled finding information one was looking for," 4 items were formulated (Cronbach alpha = .75). "Recommendation to others and future own use" was measured with 2 items (Cronbach alpha = .82). "Opinion about the number of search options" and

"overall satisfaction with the search facility" were measured with 1 item each.

To measure satisfaction with the stories retrieved, 18 items were formulated (partially based on [16,17]). "Opinion about the stories retrieved" was measured with 6 items (Cronbach alpha = .71). To measure "opinion about the list of stories displayed after a search" 4 items were formulated (Cronbach alpha = .76). "The extent to which the stories retrieved covered one's information need" was measured with 4 items (Cronbach alpha = .82) and "recommendation to others and future own reading" with 2 items (Cronbach alpha = .77). "Opinion about the number of stories retrieved" and "overall satisfaction with the stories retrieved" were measured with 1 item each.

"The stories' impact on coping with breast cancer" was measured with 6 items (Cronbach alpha = .85) which were based on an extensive literature on coping [18-21]. Two of the items were formulated to measure problem-focused coping ("By reading the stories I have learnt things" and "By reading the stories I know what to do"), another two items were formulated to measure emotion-focused coping ("By reading the stories I am more able to understand my feelings" and "By reading the stories I can see that certain emotions are part of learning to live with breast cancer"), one item was formulated to measure reappraisal ("By reading the stories I view things in a different way") and one item was formulated to measure social comparison ("By reading the stories I see that others have experienced the same things"). The existing validated coping scales, such as the Ways of Coping checklist [18,19] and the COPE inventory [21], could not be used because they were too general for our research question.

Technical Aspects

We tested the usability and technical functionality of the study website, including the final questionnaire, multiple times, and we solved all appearing errors. During participants' search processes, log data recorded how long participants surfed on the study website, how many searches they performed, how many stories they accessed, and how long the text of the stories was displayed on the screen. In the control group, clicking on a folder (A to Z) was regarded as performing a search, and subsequently clicking on a name was seen as accessing a story. Also, the time participants needed to fill in the final questionnaire was recorded.

Participants who were searching for or reading the stories were reminded to fill in the final questionnaire by a yellow figure on the left side of the screen with the text "Do not forget to complete the questionnaire," which was highlighted every five minutes. Adaptive questioning was used to reduce the number and complexity of the questions. Questions were not randomized or alternated. The final questionnaire was distributed over five pages in the following sequence:(1) the search process, (2) the stories retrieved, including the stories' impact on coping with breast cancer, (3) use of the Internet and the Amazones website, (4) disease characteristics, and (5) demographic characteristics. When participants clicked on the "Next" button at the end of a page, JavaScript was used to check for completeness. Unanswered questions were highlighted, and participants were asked to answer these. Yet, answering was not enforced, since by clicking on the "Next" button again, the next page was reached. Participants were not able to review and change their answers in previous parts to prevent a possible influence of questions asked later in the questionnaire.

Log data and questionnaire responses were saved automatically in an SQL database. In preparation for data analysis, sessions from the same IP address with a time interval of less than 20 minutes were merged, and those with a time interval of greater than 20 minutes were kept as two separate sessions. We assumed that in the former case the sessions were from the same participant and, in the latter, from different participants. Applying the first rule resulted in 23 merged sessions; the latter rule was applied to 6 pairs of sessions. Merging was possible because in all cases the questionnaire was filled out only once. Participants were only distinguished by IP address. A particular IP address was always assigned to the same intervention group. This was done to prevent women from participating multiple times when trying to get in another study group.

Data Analysis

The data were imported into SPSS version 16.0. Differences in the log data between questionnaire completers and noncompleters were assessed using Mann-Whitney tests. The noncompleters were excluded from further analyses, since no questionnaire responses for this group were available. For the completers, there was no time frame for filling in the questionnaire. Differences between the four groups in baseline characteristics were assessed using Chi square tests, 1-way ANOVA, or Kruskal-Wallis tests (depending on variable type and skewness).

Kruskal-Wallis tests were performed to assess differences between the four groups in search behaviour (ie, the log data). Significant differences were examined further by performing post hoc tests. We chose to use Mann-Whitney tests with a Bonferroni correction, and as as the critical level of significance we used .05/6 = .008 because with four groups six comparisons were performed.

For each construct of the three outcome measures (satisfaction with the search process, the stories retrieved, and the stories' impact on coping with breast cancer) a mean total score was calculated. A higher mean indicated a higher satisfaction or impact respectively. The effects of the search facilities on the constructs of the three outcome measures were examined using ANOVA with two independent factors (search facility for story topics yes/no; search facility for writer profiles yes/no) to assess possible main and interaction effects. This analytical approach was chosen in order to examine the effects of the two search facilities both independently and in combination. *P* values above .05 were considered not significant.

Results

Participant Statistics

Informed consent was given by 353 people, of whom 182 (51.6%) completed the final questionnaire (Figure 5). No significant difference was found between the four groups in percentage questionnaire completers (χ^2_3= 3.7, *P* = .30). The

mean time that participants needed to fill in the final questionnaire was 15.3 minutes (SD = 12.7; min = 5.0, max = 138.4). In comparison with questionnaire noncompleters, questionnaire completers spent less time visiting the study website (mean = 809.1 seconds vs 928.0 seconds, $P < .001$), but completers performed more searches (mean = 2.1 vs 1.7, $P < .001$), accessed more stories (mean = 6.6 vs 3.4, $P < .001$), and their mean reading time per story was longer (mean = 92.8 seconds vs 63.5 seconds, $P < .001$).

Table 3 shows the baseline characteristics of the questionnaire completers. No significant differences between the four groups were found for demographic and disease characteristics and use of the Amazones websites. With respect to use of the Internet, the writer profile group was less familiar with accessing fellow patients' stories on the Internet.

Figure 5. Flow of study participants

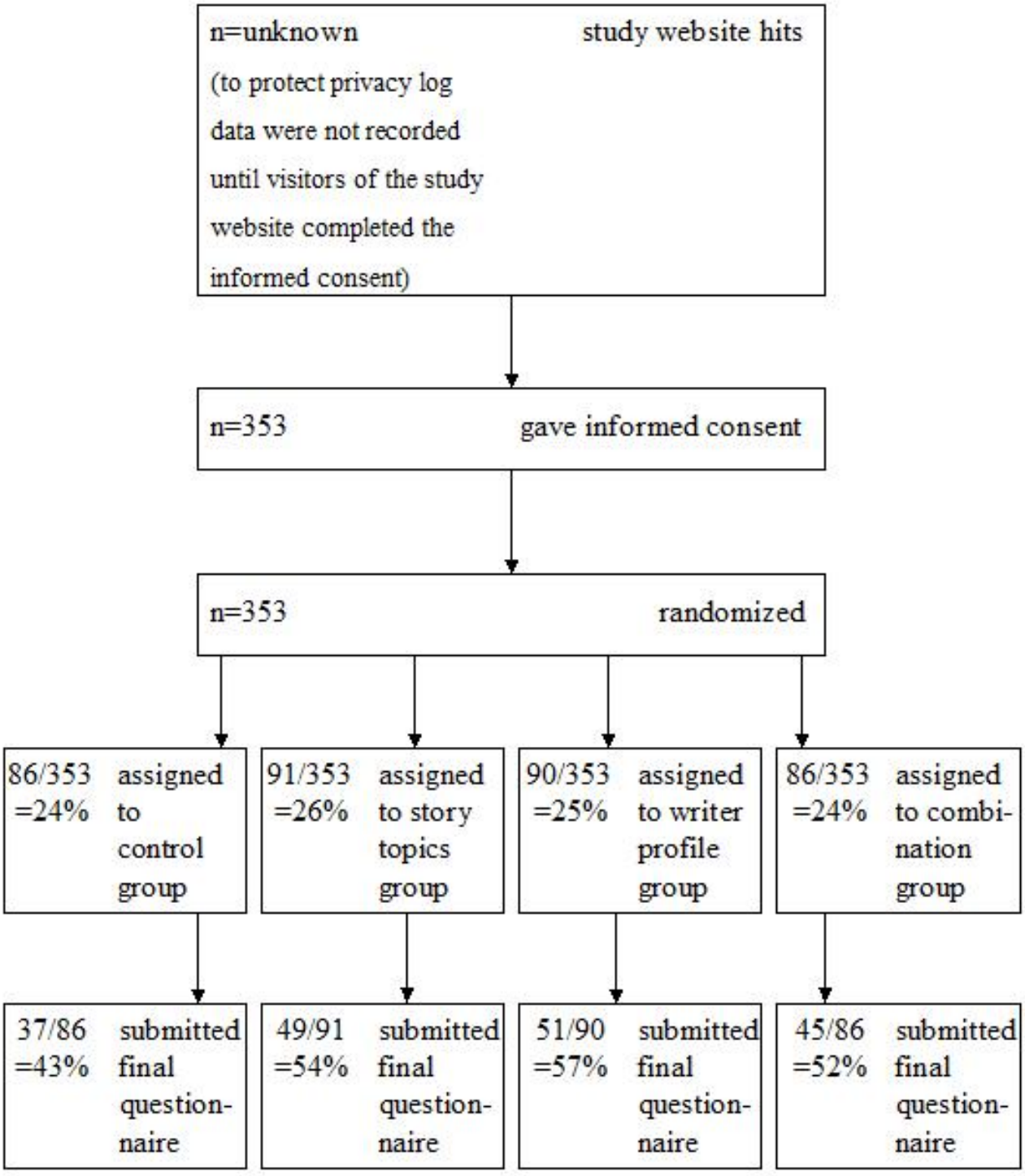


Table 3. Baseline characteristics of the four groups

		Control Group (n = 37)[a]		Story Topics Group (n = 49)[a]		Writer Profile Group (n = 51)[a]		Combination Group (n = 45)[a]		*P* value[b]
		n	%	n	%	n	%	n	%	
Demographic characteristics										
Age in years (mean, SD)		49.1 (7.5)		49.9 (8.3)		49.5 (9.4)		45.9 (9.9)		.12[c]
Married or living together	Yes	29	78.4	38	77.6	40	78.4	34	75.6	.99
	No	8	21.6	11	22.4	11	21.6	11	24.4	
Children	Yes	29	78.4	40	81.6	41	80.4	31	68.9	.45
	No	8	21.6	9	18.4	10	19.6	14	31.1	
Religious	Yes	16	43.2	19	38.8	19	37.3	12	26.7	.44
	No	21	56.8	30	61.2	32	62.7	33	73.7	
Higher professional education or university degree	Yes	11	29.7	18	36.7	23	45.1	15	33.3	.47
	No	26	70.3	31	63.3	28	54.9	30	66.7	
Employed	Yes	22	59.5	24	49.0	31	60.8	26	57.8	.64
	No	15	40.5	25	51.0	20	39.2	19	42.2	
Disease characteristics										
Time since diagnosis in months (mean, SD)		36.8 (45.8)		34.9 (41.4)		42.0 (41.8)		34.0 (37.3)		.49[d]
Diagnosed with one tumour	Yes	23	62.2	35	72.9	37	72.5	32	71.1	.69
	No	14	37.8	13	27.1	14	27.5	13	28.9	
Size of tumour	<2 cm	9	25.0	15	32.6	18	36.7	20	44.4	.32
	≥2 cm	27	75.0	31	67.4	31	63.3	25	55.6	
Cancer in axillary lymph nodes at diagnosis	Yes	15	41.7	27	57.4	24	49.0	23	51.1	.56
	No	21	58.3	20	42.6	25	51.0	22	48.9	
Metastases to other parts of the body	Yes	6	16.7	4	8.5	5	10.2	6	13.3	.68
	No	30	83.3	43	91.5	44	89.8	39	86.7	
Breast conserving surgery	Yes	13	35.1	21	42.9	18	35.3	20	44.4	.71
	No	24	64.9	28	57.1	33	64.7	25	55.6	
Mastectomy	Yes	19	51.4	24	49.0	34	66.7	30	66.7	.15
	No	18	48.6	25	51.0	17	33.3	15	33.3	
Radiation therapy	Yes	22	59.5	28	57.1	25	49.0	20	44.4	.47
	No	15	40.5	21	42.9	26	51.0	25	55.6	
Chemotherapy	Yes	24	64.9	35	71.4	31	60.8	26	57.8	.54
	No	13	35.1	14	28.6	20	39.2	19	42.2	
Hormonal therapy	Yes	21	56.8	23	46.9	23	45.1	18	40.0	.50
	No	16	43.2	26	53.1	28	54.9	27	60.0	
Cancer free	Yes	26	70.3	30	61.2	32	62.7	28	62.2	.83
	No	11	29.7	19	38.8	19	37.3	17	37.8	
Use of the Internet and the Amazones website										
Daily Internet use	Yes	31	83.8	43	87.8	47	92.2	35	77.8	.23
	No	6	16.2	6	12.2	4	7.8	10	22.2	

		Control Group (n = 37)[a]		Story Topics Group (n = 49)[a]		Writer Profile Group (n = 51)[a]		Combination Group (n = 45)[a]		*P* value[b]
		n	%	n	%	n	%	n	%	
Familiar with searching online for specific information	Yes	36	97.3	46	93.9	48	94.1	43	95.6	.88
	No	1	2.7	3	6.1	3	5.9	2	4.4	
Familiar with accessing fellow patients' stories on the Internet	Yes	30	81.1	44	89.8	33	64.7	40	88.9	.005
	No	7	18.9	5	10.2	18	35.3	5	11.1	
Visited the Amazones website at least once a month before participation	Yes	11	29.7	18	36.7	15	29.4	19	42.2	.52
	No	26	70.3	31	63.3	36	70.6	26	57.8	
"Rather well" or "well" familiar with Amazones website	Yes	7	38.9	18	52.9	15	50.0	13	38.2	.56[e]
	No	11	61.1	16	47.1	15	50.0	21	61.8	
Read half or more of the Amazones stories before	Yes	5	27.8	15	44.1	13	43.3	12	35.3	.62[e]
	No	13	72.2	19	55.9	17	56.7	22	64.7	

[a]N (%) is shown unless noted otherwise.

[b]*P* values are for chi-square tests comparing the four groups unless noted otherwise.

[c]*P* value for 1-way ANOVA test to compare the four groups with respect to age.

[d]*P* value for Kruskal-Wallis test to compare the four groups with respect to time since diagnosis.

[e]Percentages and tests based on the number of participants who had previously visited the Amazones website: control group (n = 18), story topics group (n = 34), writer profile group (n = 30), combination group (n = 34).

Search Behaviour

Table 4 shows that there were no differences between the four groups in time spent on the study website or in the number of searches performed. However, we found differences between the four groups in the number of stories accessed and in the mean reading time per participant per story. Post hoc tests (with a critical level of significance of $P = .008$ due to Bonferroni correction) showed that compared with the control group, fewer stories tended to be accessed in both the writer profile group ($P = .01$) and the combination group ($P = .02$). In addition, in the control group, the mean reading time per participant per story was shorter than in the writer profile group ($P = .007$) and tended to be shorter compared with the story topics group ($P = .02$) and the combination group ($P = .009$).

Satisfaction With the Search Process

Table 5 shows that having access to the story topics search facility resulted in a more positive opinion about the search facility (1a), in a more positive opinion about the number of search options (1b), in being better enabled to find the information one was looking for (1c), in being more inclined to recommend it to others or to use it more often themselves in future (1d), and in a higher overall satisfaction with the search facility (1e), compared with not having access to this search facility (all comparisons were significant at $P < .05$). Having access to the writer profile search facility compared with not having access to this search facility resulted in a significantly more positive opinion about the search facility (1a), and in a significantly higher overall satisfaction score (1e).

An interaction effect was found for the overall satisfaction score (1e). When participants could search using the story topics, they were satisfied with this search facility regardless of whether (mean = 7.3, SD = 1.5) or not (mean = 7.2, SD = 1.4) they could also search with the writer profile. The effect of having access to the writer profile search facility when also having access to the story topics search facility was not significant ($P = .90$). However, when participants could not use the story topics to search the stories, they were more satisfied with having access to the writer profile as a search facility (mean = 6.8, SD = 1.6) compared with not having access to any search facility (mean = 5.7, SD = 2.3). The effect on satisfaction of having access to the writer profile search facility when not having access to the story topics search facility was significant ($P = .009$).

Satisfaction With the Stories Retrieved

Having access to the story topics search facility resulted in a more positive opinion about the list of stories displayed after a search (2c), a greater extent to which the stories retrieved covered one's information need (2d), and a higher overall satisfaction score with the stories retrieved (2f) compared with not having access to this search facility (Table 5). Having access to the writer profile search facility compared with not having access to this search facility resulted in a more positive opinion about the list of stories displayed after a search (2c).

There were no interaction effects observed in satisfaction with the stories retrieved.

The Stories' Impact on Coping With Breast Cancer

Table 5 shows that the stories retrieved using the story topics search facility had a greater impact on coping with breast cancer (3a). When analysing each of the six coping items individually, we observed that having access to the story topics search facility

resulted in a significantly higher score for having learned things (3a.1).

Table 4. Comparison of the four groups for the search behaviour measures recorded by the log data

		Control Group (n = 37)	Story Topics Group (n = 49)	Writer Profile Group (n = 51)	Combination Group (n = 45)	*P* value[a]
Time spent on the study website in seconds	mean (SD)	754.00 (966.33)	984.55 (1278.94)	595.39 (630.04)	905.49 (1054.71)	.45
	median	496.00	634.00	389.00	636.00	
Number of searches	mean (SD)	3.89 (4.71)	1.88 (2.32)	1.53 (1.59)	1.69 (1.58)	.07
	median	2.50	1.23	1.24	1.28	
Number of stories accessed	mean (SD)	13.19 (18.98)	6.73 (6.36)	4.18 (4.53)	3.93 (3.61)	.01
	median	5.67	5.56	3.00	3.29	
Reading time per participant per story in seconds	mean (SD)	49.24 (54.65)	94.16 (94.90)	99.72 (128.27)	119.26 (113.00)	.02
	median	28.11	71.33	67.00	89.00	

[a]*P* value for Kruskal-Wallis tests comparing the four groups with respect to the four search behaviour measures.

Table 5. Means (SD) of the constructs of the three outcome measures asked for in the final questionnaire by search factor

	Story Topics[a]			Writer Profile[a]			Interaction[a]
	Yes (n = 94)	No (n = 88)	*P* value[b]	Yes (n = 96)	No (n = 86)	*P* value[c]	*P* value[d]
1. Satisfaction with the search process							
a. opinion about the search facility (range 1-5)	4.0 (0.7)	3.6 (1.1)	.001	3.9 (0.9)	3.6 (1.0)	.005	.21
b. opinion about the number of search options (range 1-3)[e]	2.3 (0.6)	2.1 (0.8)	.04	2.3 (0.7)	2.2 (0.8)	.29	.23
c. the extent to which search options enable finding information one was looking for (range 1-5)	3.3 (1.0)	2.8 (1.0)	.001	3.1 (1.0)	3.0 (1.0)	.27	.59
d. recommendation to others and future own use (range 1-5)	4.1 (1.0)	3.5 (1.2)	< .001	3.9 (1.1)	3.8 (1.2)	.29	.13
e. overall satisfaction with the search facility (range 1-10)	7.3 (1.4)	6.3 (2.0)	< .001	7.1 (1.6)	6.5 (2.0)	.01	.03
2. Satisfaction with (the information in) the stories retrieved							
a. opinion about the stories retrieved (range 1-5)	3.5 (0.6)	3.4 (0.7)	.54	3.5 (0.6)	3.4 (0.7)	.36	.18
b. opinion about the number of stories retrieved (range 1-3)[e]	2.3 (0.7)	2.1 (0.7)	.27	2.1 (0.7)	2.3 (0.7)	.18	.17
c. opinion about the list of stories displayed after a search (range 1-5)	3.6 (0.9)	3.2 (1.2)	.001	3.8 (0.9)	2.9 (1.1)	< .001	.06
d. the extent to which the stories retrieved covered one's information need (range 1-5)	3.0 (1.0)	2.6 (1.0)	.02	2.7 (1.1)	2.9 (1.0)	.56	.91
e. recommendation to others and future own reading (range 1-5)	4.1 (1.0)	3.8 (1.0)	.08	4.0 (1.0)	4.0 (1.0)	.67	.71
f. overall satisfaction with the stories retrieved (range 1-10)	7.1 (1.5)	6.4 (1.7)	.002	6.7 (1.7)	6.9 (1.6)	.80	.35
3. The stories' impact on coping with breast cancer							
a. the stories' impact on coping with breast cancer (range 1-5)	3.2 (0.9)	2.9 (1.0)	.02	3.0 (1.0)	3.1 (1.0)	.53	.71
a.1. By reading the stories I have learned things (range 1-5)	3.0 (1.4)	2.5 (1.3)	.007	2.6 (1.4)	2.9 (1.3)	.24	.53
a.2. By reading the stories I know what to do (range 1-5)	2.7 (1.2)	2.4 (1.2)	.14	2.5 (1.3)	2.6 (1.1)	.71	.70
a.3. By reading the stories I am more able to understand my feelings (range 1-5)	3.0 (1.3)	2.6 (1.4)	.07	2.7 (1.4)	2.9 (1.3)	.70	.37
a.4. By reading the stories I can see that certain emotions are part of learning to live with breast cancer (range 1-5)	3.8 (1.2)	3.5 (1.5)	.17	3.7 (1.4)	3.6 (1.3)	.82	.76
a.5. By reading the stories I view things in a different way (range 1-5)	2.7 (1.3)	2.4 (1.3)	.16	2.4 (1.3)	2.7 (1.3)	.12	.16
a.6. By reading the stories I see that others have experienced the same things (range 1-5)	4.1 (1.0)	3.9 (1.3)	.16	4.1 (1.1)	3.9 (1.1)	.28	.31

[a] ANOVA with two independent factors (search facility for story topics yes/no; search facility for writer profiles yes/no). Higher means indicate better outcomes.

[b] *P* value for possible main effect of story topics search

[c] *P* value for possible main effect of writer profile search

[d] *P* value for possible interaction effect between story topics search and writer profile search

[e] Asked on a 5-point scale but for analysis recoded into 3-points (see also Multimedia Appendix 3)

Discussion

Principal Findings

To our knowledge, this study is the first randomized controlled experiment with a 2x2 factorial design that examined search facilities for accessing online patient stories. We observed that the story topics search factor had a strong impact on patient satisfaction and search success: participants were the most satisfied with this search facility and the stories retrieved. Also, the stories retrieved had a greater impact on coping with breast cancer. The effect of the writer profile search factor was limited. This search facility resulted only in a few effects, predominantly on satisfaction with the search process. The two search factors combined generally had no amplified effect on patient satisfaction or search success as we only had one significant interaction.

These findings are contrary to our expectation, which was that the combined search facilities (the interaction) would outperform a single search facility because this combination is more complete and differentiated resulting in greater opportunities to find a relevant story. Apparently, this quantity argument seems to be less important than the type of the search facility (quality). In line with our expectation was that a single search facility was an improvement compared with the alphabetically listed stories in the control group.

Participants in the three search facility groups accessed fewer stories and read longer per accessed story compared with the control group. An explanation for this might be that the stories retrieved in the search facility groups were more relevant to the participants. A search facility probably increases the proportion of the documents retrieved relevant to the user's information need [22].

The story topics search facility resulted not only in participants being more satisfied with the search process, but also in participants retrieving stories that better covered their information needs and retrieving stories from which they learned more. Patients might use online stories predominantly for information, and, therefore, the topics described in the stories might be more important for them than the writer's profile. Patients' profiles might be more important when seeking face-to-face contact. This difference between seeking information and seeking contact has also been noted by Bennenbroek et al [23] in their research on social comparison.

Our observation that the writer profile search facility compared with not having this search facility resulted in a more positive opinion about the search facility and in a higher overall satisfaction with the search facility is in line with the results of Rozmovits and Ziebland [1]. They found in interviews that patients positively evaluated the ability to select other patients of a particular age, stage of illness, or patients who were long-term survivors or who had opted for similar treatment. However, although our study also showed that participants were more satisfied with this search facility, they were not more satisfied with the stories retrieved using this facility.

Limitations of the Present Study

A considerable number of participants performed searches but did not complete the questionnaire. Compared with completers, noncompleters spent more time on the study website while they performed fewer searches, accessed fewer stories, and spent less time reading per story. Noncompleters might not have been sure about how to use the search facilities, or they might not have been as interested. Yet, we could not empirically evaluate these hypotheses nor perform any statistical analyses, since we had no further information about noncompleters. The number of completers and noncompleters was evenly distributed over the four study groups. Therefore, we believe that potential bias equally affected all four groups. In addition, the direction of the bias is probably twofold: dissatisfied participants might have stopped or they might have completed the questionnaire to express their annoyance.

More than half of the participants who completed the questionnaire (63.7%) had previously visited the Amazones website. This could introduce bias because participants familiar with the original disclosure of stories might be especially satisfied with the new search facilities. However, frequency of visiting the Amazones site, knowing the site "rather well" or "well," and the number of Amazones stories read before, were all evenly distributed over the four study groups. Therefore, we do not think this previous experience with the Amazones website affected the results.

Since the experiment was conducted completely online [24], we cannot verify that all participants indeed had (or had had) breast cancer. However, we targeted this group for recruitment and asked relevant questions before randomization. We assume that all participants were sincere, because overall they spent 15 minutes filling in the final questionnaire. This suggests that participants were interested in the subject matter.

A limitation of the design was a possible confounding between type of search facility (story topics, writer profile) and number of search options (17 topics, 6 personal characteristics). The story topics search might have been more appreciated because it was more extensive than the writer profile search. However, an argument against this reasoning is that the most extensive search facility (ie, story topics in combination with writer profile) was not the most favourite.

In addition, one could question the content of the search facilities. Were the most appropriate topics and personal characteristics included in the facilities? Yet, the topics and characteristics we used were chosen based on other websites containing breast cancer stories [11-13] and other studies in this field [1,9,10].

A final limitation is that participants may have been annoyed when stories were presented that did not exactly match their search objectives. However, a search resulting in no stories could also be a cause of annoyance. This is why we chose to present at least ten stories after each search. In order to ease interpretation of the resulting list of stories, weights (as pictured in the form of pink ribbons) were used to indicate to what extent a story matched with the search objective.

Conclusions and Practical Implications

Earlier studies have shown that patients can benefit from stories of other patients, and that the Internet is an important source of these stories. Our current study suggests that a story topics search facility would be most helpful to patients. With a story topics search facility, participants were better enabled to find the information they were looking for. Also, they retrieved stories that more closely covered their information needs and they learned more from the stories retrieved.

Thus, patient organisations or website developers that offer patient stories on their websites can best provide access to them using a story topics search facility. However, constructing such a search facility is very time consuming and labour intensive since stories have to be coded for content. An efficient method might be to use a system analogous to social bookmarking/tagging [25] in which story readers assign keywords or tags to the stories, and the keywords or tags that are most often assigned are seen as most important in describing the content. Another possibility is to construct a list of items from which writers can compose descriptions of their stories. Finally, stories could also be classified by automatic full text indexing or clustering. This will be the subject of our next study.

Acknowledgments

The authors wish to thank the participants for their time and interest. We also wish to thank the organizations that disseminated the call for participation on their websites or in their waiting rooms.

The paper is part of the Narrator project addressing the problem of storage and accurate retrieval of illness narratives. This project is conducted under the umbrella program ToKeN2000 and is funded by the Netherlands Organization for Scientific Research (NWO).

Conflicts of Interest

None declared

Multimedia Appendix 1

Research proposal (in Dutch) as presented to the ethical committee before the start of the study

[PDF file (Adobe PDF),291 KB - jmir_v12i1e7_app1.pdf]

Multimedia Appendix 2

Screenshots illustrating the search page and the list of retrieved stories for each of the 4 search conditions (ie, control, story topics, writer profile, and combination condition)

[PDF file (Adobe PDF),943 KB - jmir_v12i1e7_app2.pdf]

Multimedia Appendix 3

An overview of the constructs and items belonging to the three main outcome measures

[PDF file (Adobe PDF),70 KB - jmir_v12i1e7_app3.pdf]

References

1. Rozmovits L, Ziebland S. What do patients with prostate or breast cancer want from an Internet site? A qualitative study of information needs. Patient Educ Couns 2004 Apr;53(1):57-64. [Medline: 15062905] [doi: 10.1016/S0738-3991(03)00116-2]
2. Ziebland S, Chapple A, Dumelow C, Evans J, Prinjha S, Rozmovits L. How the internet affects patients' experience of cancer: a qualitative study. BMJ 2004 Mar 6;328(7439):564 [FREE Full text] [Medline: 15001506] [doi: 10.1136/bmj.328.7439.564]
3. McTavish FM, Gustafson DH, Owens BH, Hawkins RP, Pingree S, Wise M, et al. CHESS (Comprehensive Health Enhancement Support System): an interactive computer system for women with breast cancer piloted with an underserved population. J Ambul Care Manage 1995 Jul;18(3):35-41. [Medline: 10143478]
4. Herxheimer A, McPherson A, Miller R, Shepperd S, Yaphe J, Ziebland S. Database of patients' experiences (DIPEx): a multi-media approach to sharing experiences and information. Lancet 2000 Apr 29;355(9214):1540-1543. [Medline: 10801187] [doi: 10.1016/S0140-6736(00)02174-7]
5. Pitts V. Illness and Internet empowerment: writing and reading breast cancer in cyberspace. Health (London) 2004 Jan;8(1):33-59. [Medline: 15018717] [doi: 10.1177/1363459304038794]
6. Hardey M. 'The story of my illness': personal accounts of illness on the Internet. Health (London) 2002;6:31-46. [doi: 10.1177/136345930200600103]
7. McLellan MF. The Electronic Narrative of Illness [dissertation]. Galveston: University of Texas Medical Branch; 1997.

8. Wise M, Han JY, Shaw B, McTavish F, Gustafson DH. Effects of using online narrative and didactic information on healthcare participation for breast cancer patients. Patient Educ Couns 2008 Mar;70(3):348-356 [FREE Full text] [Medline: 18201859] [doi: 10.1016/j.pec.2007.11.009]
9. Overberg RI, Alpay LL, Verhoef J, Zwetsloot-Schonk JHM. Illness stories on the internet: what do breast cancer patients want at the end of treatment? Psychooncology 2007 Oct;16(10):937-944. [Medline: 17274107] [doi: 10.1002/pon.1157]
10. Williamson K. Where one size does not fit all: understanding the needs of potential users of a portal to breast cancer knowledge online. J Health Commun 2005 Sep;10(6):567-580. [Medline: 16203634] [doi: 10.1080/10810730500228961]
11. Healthtalkonline (formerly DIPEx). Health Experiences Research Group URL: http://www.healthtalkonline.org/Cancer/Breast_Cancer [WebCite Cache ID 5eAfXVpWN]
12. BreastCancerStories.com. URL: http://www.BreastCancerStories.com [WebCite Cache ID 5eAg5Mvy9]
13. KWF Kankerbestrijding (Dutch Cancer Society). URL: http://www.kankerbestrijding.nl [WebCite Cache ID 5eAgM1yAc]
14. De Amazones, jonge vrouwen met borstkanker (The Amazones, young women with breast cancer). URL: http://www.de-amazones.nl [WebCite Cache ID 5eAgT7f7P]
15. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [Medline: 15471760] [doi: 10.2196/jmir.6.3.e34]
16. Spink A. A user-centered approach to evaluating human interaction with Web search engines: an exploratory study. Info Process Manag 2002;38(3):401-426. [doi: 10.1016/S0306-4573(01)00036-X]
17. Zeng QT, Kogan S, Plovnick RM, Crowell J, Lacroix EM, Greenes RA. Positive attitudes and failed queries: an exploration of the conundrums of consumer health information retrieval. Int J Med Inform 2004 Feb;73(1):45-55. [Medline: 15036078] [doi: 10.1016/j.ijmedinf.2003.12.015]
18. Folkman S, Lazarus RS. An analysis of coping in a middle-aged community sample. J Health Soc Behav 1980 Sep;21(3):219-239. [Medline: 7410799] [doi: 10.2307/2136617]
19. Folkman S, Lazarus RS. If it changes it must be a process: study of emotion and coping during three stages of a college examination. J Pers Soc Psychol 1985 Jan;48(1):150-170. [Medline: 2980281] [doi: 10.1037/0022-3514.48.1.150]
20. McCrae RR, Costa PT. Personality, coping, and coping effectiveness in an adult sample. Journal of Personality 1986;54:385-405. [doi: 10.1111/j.1467-6494.1986.tb00401.x]
21. Carver CS, Scheier MF, Weintraub JK. Assessing coping strategies: a theoretically based approach. J Pers Soc Psychol 1989 Feb;56(2):267-283. [Medline: 2926629] [doi: 10.1037/0022-3514.56.2.267]
22. Van Rijsbergen CJ. Information Retrieval. 2nd edition. London, United Kingdom: Butterworths; 1979.
23. Bennenbroek FTC, Buunk BP, van der Zee KI, Grol B. Social comparison and patient information: what do cancer patients want? Patient Educ Couns 2002 May;47(1):5-12. [Medline: 12023095] [doi: 10.1016/S0738-3991(02)00018-6]
24. Nosek BA, Banaji MR, Greenwald AG. E-Research: Ethics, security, design, and control in psychological research on the Internet. J Soc Issues 2002;58:161-176. [doi: 10.1111/1540-4560.00254]
25. Kamel Boulos MN, Wheeler S. The emerging Web 2.0 social software: an enabling suite of sociable technologies in health and health care education. Health Info Libr J 2007;24:2-23. [doi: 10.1111/j.1471-1842.2007.00701.x]

Edited by G Eysenbach; submitted 02.02.09; peer-reviewed by J Bender, J Fogel; comments to author 27.02.09; revised version received 31.12.09; accepted 20.01.10; published 09.03.10.

Please cite as:
Overberg R, Otten W, Man AD, Toussaint P, Westenbrink J, Zwetsloot-Schonk B
How Breast Cancer Patients Want to Search for and Retrieve Information From Stories of Other Patients on the Internet: an Online Randomized Controlled Experiment
J Med Internet Res 2010;12(1):e7
URL: http://www.jmir.org/2010/1/e7/
doi:10.2196/jmir.1215
PMID:20215101

Original Paper

Comparison of Trial Participants and Open Access Users of a Web-Based Physical Activity Intervention Regarding Adherence, Attrition, and Repeated Participation

Miriam Wanner[1,2], PhD; Eva Martin-Diener[1], MPH; Georg Bauer[3,4], MD, DrPH; Charlotte Braun-Fahrländer[2], MD; Brian W Martin[3], MD, MPH

[1]Swiss Federal Institute of Sport Magglingen, Magglingen, Switzerland

[2]Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland

[3]Institute of Social and Preventive Medicine, University of Zurich, Zurich, Switzerland

[4]Division of Public and Organizational Health, University and ETH Zurich, Zurich, Switzerland

Corresponding Author:
Miriam Wanner, PhD
Swiss Federal Institute of Sport Magglingen
Brésil, Alpenstrasse 21+23
CH-2532 Magglingen
Switzerland
Phone: +41 32 327 6138
Fax: +41 32 327 6404
Email: miriam.wanner@alumni.ethz.ch

Abstract

Background: Web-based interventions are popular for promoting healthy lifestyles such as physical activity. However, little is known about user characteristics, adherence, attrition, and predictors of repeated participation on open access physical activity websites.

Objective: The focus of this study was Active-online, a Web-based individually tailored physical activity intervention. The aims were (1) to assess and compare user characteristics and adherence to the website (a) in the open access context over time from 2003 to 2009, and (b) between trial participants and open access users; and (2) to analyze attrition and predictors of repeated use among participants in a randomized controlled trial compared with registered open access users.

Methods: Data routinely recorded in the Active-online user database were used. Adherence was defined as: the number of pages viewed, the proportion of visits during which a tailored module was begun, the proportion of visits during which tailored feedback was received, and the time spent in the tailored modules. Adherence was analyzed according to six one-year periods (2003-2009) and according to the context (trial or open access) based on first visits and longest visits. Attrition and predictors of repeated participation were compared between trial participants and open access users.

Results: The number of recorded visits per year on Active-online decreased from 42,626 in 2003-2004 to 8343 in 2008-2009 (each of six one-year time periods ran from April 23 to April 22 of the following year). The mean age of users was between 38.4 and 43.1 years in all time periods and both contexts. The proportion of women increased from 49.5% in 2003-2004 to 61.3% in 2008-2009 ($P< .001$). There were differences but no consistent time trends in adherence to Active-online. The mean age of trial participants was 43.1 years, and 74.9% were women. Comparing contexts, adherence was highest for registered open access users. For open access users, adherence was similar during the first and the longest visits; for trial participants, adherence was lower during the first visits and higher during the longest visits. Of registered open access users and trial participants, 25.8% and 67.3% respectively visited Active-online repeatedly ($P< .001$). Predictors of repeated use were male sex (odds ratio [OR] = 1.2, 95% confidence interval [CI] = 1.04-1.38) and increasing age category in registered open access users, and age 46-60 versus < 30 years (OR = 3.04, 95% CI = 1.25-7.38) and Swiss nationality ($OR_{nonSwiss}$= 0.64, 95% CI = 0.41-1.00) in trial participants. Despite reminder emails, attrition was much higher in registered open access users compared with trial participants, with a median lifetime website usage of 0 days in open access users and 290 days in trial participants.

Conclusions: Adherence, patterns of use, attrition, and repeated participation differed between trial participants and open access users. Reminder emails to encourage repeated participation were effective for trial participants but not for registered open access users. These issues are important when interpreting results of randomized controlled effectiveness trials.

(J Med Internet Res 2010;12(1):e3) doi:10.2196/jmir.1361

KEYWORDS

Internet; individually tailored intervention; user characteristics; time trends

Introduction

In recent years,Web-based interventions targeting health issues such as nutrition [1], smoking [2], physical activity [3,4], or multiple health behaviors [5] have become popular. These interventions have several advantages, such as interactive designs, the possibility of tailoring information to individual users, the potential to reach large audiences at relatively low costs, and the ease with which users can get involved, that is, people can use the intervention at home.

To maximize effectiveness, it is important for developers of such interventions to know more about user characteristics, adherence (the extent to which individuals use the content of the Internet intervention) [6], nonusage attrition (whether individuals discontinue use of an Internet intervention) [7], and predictors of repeated Internet intervention use. Studies reporting these issues have done so mostly in the context of randomized controlled trials (RCT) [8-13] or in other controlled study settings [14]. However, the use of the intervention in a trial context may not reflect the use of the intervention in an open access context. This is an important issue, as program effectiveness is likely to depend on adherence to and use of websites. If adherence is higher among trial participants (eg, due to higher motivation among trial participants and the efforts of the study staff to increase adherence), effectiveness may be overestimated; if adherence is lower among trial participants (eg, due to a higher burden of additional data assessments), effectiveness may be underestimated.

Few studies have described use and users of open access websites in the domain of smoking [15,16], mental health [17,18], and drinking [19]. Even fewer studies have done so in the domain of physical activity. To our knowledge, there is only one study that has described rates and determinants of repeated participation in an open access Web-based program aimed at healthy lifestyles that has emphasized healthy body weight and physical activity [20].

In Switzerland, a Web-based tailored physical activity intervention (Active-online) [21] developed between 1999 and 2003, has been freely available as an open access program since 2003. Continuous data collection pertaining to website visitors provides an opportunity to analyze user characteristics and patterns of intervention use and adherence over time.The effectiveness of the intervention was assessed in a Web-based RCT in 2006-2007 [22] after the study design had been tested in a feasibility study in 2003 [23].

The aims of the present study were: (1) to assess and compare user characteristics and adherence to the website (a) in an open access context over time from 2003 to 2009, and (b) between participants in an RCT and open access users (all open access users and the subgroup of registered open access users only), and (2) to analyze attrition and potential predictors of repeated use of the website in trial participants compared with registered open access users.

Methods

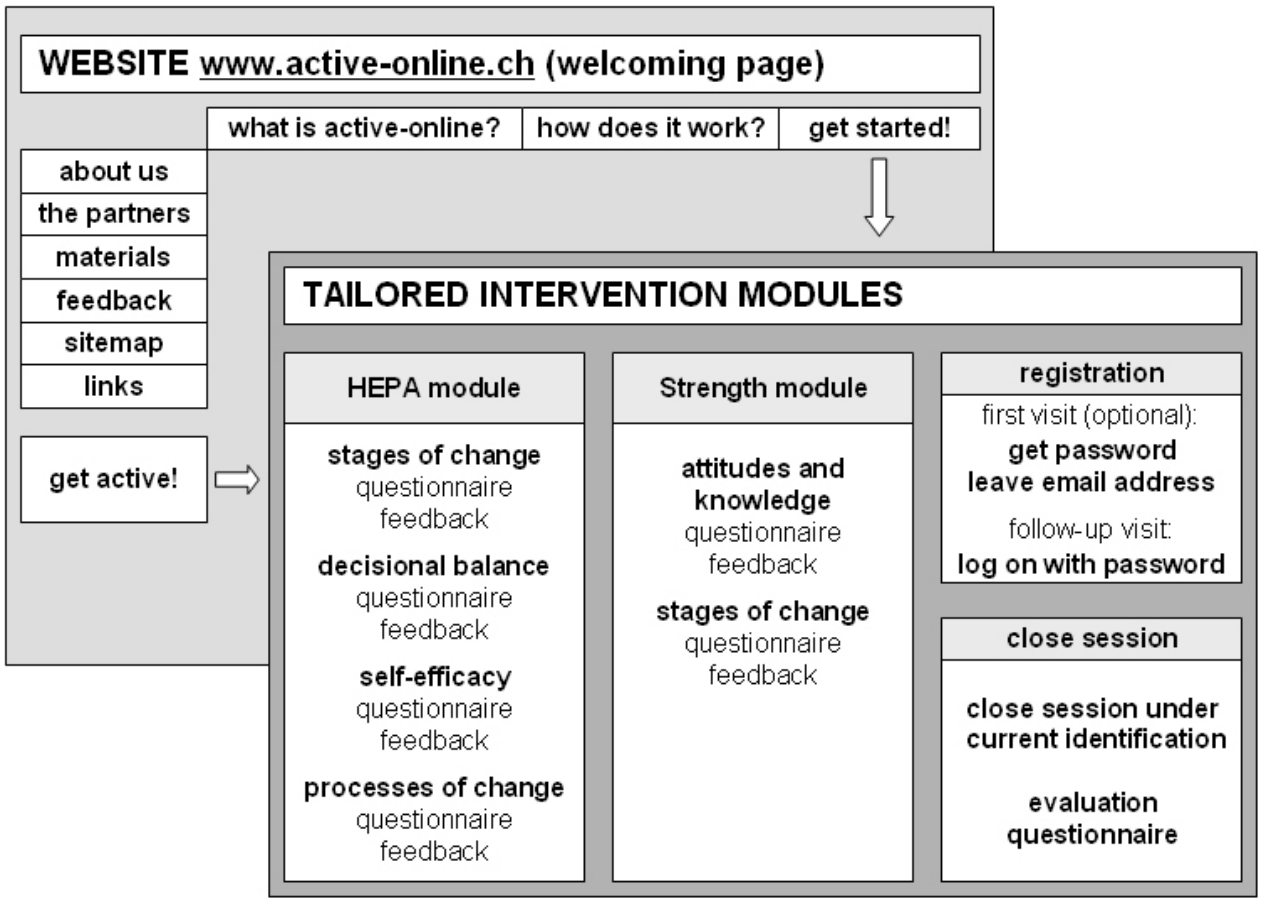
Intervention Program

Active-online is an individually tailored program to promote physical activity targeting adults aged 30 to 60 years. Active-online is freely available on the Internet in the three main languages of Switzerland: German, French, and Italian. At the start of the program, users find a language selection page followed by a welcoming page that explains the program and provides additional information and motivational material (see Figure 1 and Figure 2). At the beginning of the tailored intervention, a new window opens where two modules are offered: one module on everyday health-enhancing physical activity (HEPA) and endurance training (the HEPA module), and one module on strength and flexibility training (the strength module) (Figure 2). Visits are recorded in the database as soon as a new window is opened (after the welcoming page) where one of the two tailored modules can be selected.

Users may visit Active-online without registering, or they may register. Registration is very brief and involves leaving an email address. Registered users receive a password which allows them to revisit the website and follow changes in their physical activity behavior. Registered users receive reminder emails that contain a link to revisit the website after 2, 4, and 7 months.

Figure 1. Screenshot of the welcoming page of Active-online

Figure 2. Structure of Active-online and the tailored intervention modules



The HEPA module offers a maximum of four individually tailored feedbacks based on the transtheoretical model of behavior change [24]. Stages of change are assessed according to a seven-stage concept focusing on current behavior (moderate- and vigorous-intensity activities) as well as on intention to change [25]. The module on strength and flexibility

training offers a maximum of two tailored feedbacks. Questionnaires preceding each feedback include between 5 and 23 questions and can be completed in a few minutes each. It is unlikely that visitors complete all parts of both tailored modules during one visit, thus repeated visits are encouraged (with reminder emails to registered users). Moreover, repeated visitors who answer the questions obtain individually tailored feedback on changes since their last visit.

Randomized Controlled Effectiveness Trial

The RCT is described in detail elsewhere [22]. Briefly, participants were recruited in 2006 by advertisements in magazines, newspapers, and on websites. The advertisements asked for volunteers to participate in a Web-based physical activity study. After completing an online baseline questionnaire with items on demographics, general health, and physical activity behavior, participants assigned to the intervention group were forwarded to the open access program, Active-online, and directed to use the intervention. (Participants assigned to the control group were forwarded to a simple nontailored website that contained general information on physical activity and health.) To replicate the conditions of open access use, participants assigned to the intervention group had access to the general instructions on the website; they did not receive additional instructions in how to use the intervention. All contacts were by email. Trial participants received reminder emails to revisit the Active-online website at 9, 10, and 11 months after baseline. (Participants assigned to the control group received no reminder emails.) Follow-up assessments as part of the requirements of the RCT took place at 6 weeks, 6 months, and 13 months. While registration on Active-online was not compulsory for open access users, trial participants were automatically registered in order to analyze their intervention use.

Time Periods of Intervention Use, Data Collection, and Variables Included in the Analyses

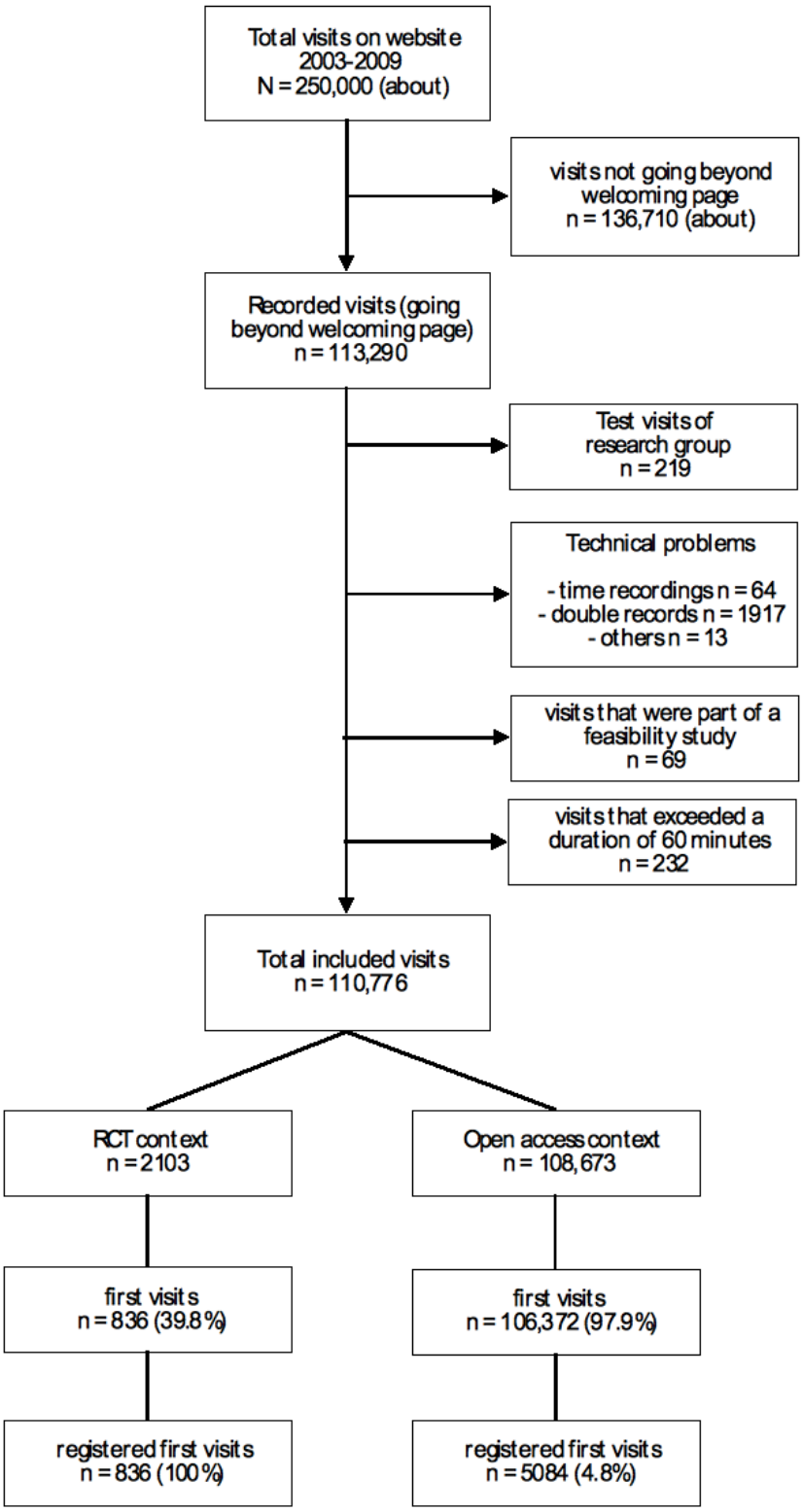
In April 2003, Active-online was officially launched with a promotional event. Data from the open access period were included for six one-year time periods through April 2009. Each one-year time period started on April 23 and ended on April 22 of the following year. Data from the effectiveness trial were included from May 1, 2006, through September 30, 2007.

The total number of visits on Active-online (including visits on the welcoming page) was available from the Internet provider. However, the absolute numbers were difficult to interpret because these depended on the software used to assess them and whether visits by web crawlers could be identified, for example. For both open access users and trial participants, visits on Active-online were captured in the Active-online user database as soon as the new browser window for the selection of one of the two tailored modules was opened (see Figure 2), whereas visits that did not go beyond the welcoming page were not recorded.

For each visit that was captured in the Active-online user database, starting time and date, finishing time and date, number of pages viewed, and time spent within the tailored modules were recorded in addition to responses to the questionnaires preceding each tailored feedback.

During the study period, between April 23, 2003, and April 22, 2009, more than 250,000 visits were counted on Active-online (including visits not going beyond the welcoming page and excluding visits by web crawlers). For the present study, only those visits that went beyond the welcoming page and were recorded in the Active-online user database during the study period were included. These numbered 113,290. Figure 3 shows the inclusion and exclusion of visits that served as the basis for the analyses in this study. We excluded 219 visits that were obvious test visits of the research group and another 64 visits because the start times and end times did not correspond. Furthermore, 232 visits were excluded because they exceeded 60 minutes as we assumed these visits were not properly terminated. Users who received all tailored feedbacks, read them online, and used the option to download further materials could easily have spent half an hour or more on Active-online.The rationale to use a cutoff of 60 minutes was to allow for potential distractions during a visit and to make sure that no serious visits were excluded. Another 13 visits were excluded due to other recording problems, and 1917 records were deleted because they represented double visits of the same session. Finally, 69 records of individuals who had participated in the feasibility study [23] were dropped. Thus, 110,776 visits were included, of which 107,208 (96.8%) were recorded as first visits.

Figure 3. Inclusion and exclusion of records used in the analyses



Of the 110,776 visits, 108,673 were recorded in the open access context and 2103 during the RCT. For the analysis of time trends, open access visits were stratified according to the six one-year periods between 2003 and 2009. The number of recorded and first visits and the proportion of first visits resulting in registration are displayed in Table 1 for each time period and for the RCT. Of all open access visits, 2.1% were recorded as repeated visits. The corresponding proportion was 60.2% during the RCT.

Table 1. Recorded website visits in the open access context (according to the six time periods from 2003-2009) and during the RCT

Open Access Program Use	Number of Visits Recorded in Database	Number of First Visits (% of Recorded Visits)	Number of First Visits Resulting in Registration (% of First Visits)
Time period (from April 23 to April 22 of the following year)			
2003-2004	42,626	41,699 (97.8%)	2263 (5.4%)
2004-2005	25,392	25,026 (98.6%)	784 (3.1%)
2005-2006	12,776	12,517 (98.0%)	592 (4.7%)
2006-2007	9847	9539 (96.9%)	610 (6.4%)
2007-2008	9689	9451 (97.5%)	513 (5.4%)
2008-2009	8343	8140 (97.6%)	322 (4.0%)
Total (2003-2009)	108,673	106,372 (97.9%)	5084 (4.8%)
RCT[a] (May 1, 2006 to September 30, 2007)	**2103**	**836 (39.8%)**	**836 (100%)**

[a]Trial participants were automatically registered within Active-online.

The proportion of women among Active-online users and the mean age of visitors were included in the analyses as demographic variables. The main measure of physical activity was the proportion meeting the current Swiss recommendations for health-enhancing physical activity (HEPA): 30 minutes or more of moderate intensity activities on 5 or more days per week or 20 minutes or more of vigorous intensity activities on 3 or more days per week [26]. Additional potential predictors of repeated participation were available for trial participants only. These were smoking, BMI, education, and nationality. For the comparison of repeated participation between open access users and trial participants, only registered open access users were included (n = 5084, 4.8% of all open access users) because repeated visits could only be tracked for participants who had registered.

Adherence, defined as the extent to which individuals experienced the content of the website [6], is reported as the number of pages viewed on Active-online, the proportion of visits that resulted in starting a tailored module, the proportion of visits when at least 3 minutes were spent in a tailored module (assuming that a minimum of 3 minutes is required to get involved with the intervention), the proportion of visits when at least one tailored feedback (HEPA or strength module, see Figure 2) was received, and the time spent in the tailored modules. These measures of adherence are commonly used to describe the extent to which individuals use the material on Web interventions [6,13]. Analysis of adherence was based on first visit to compare open access dissemination of the intervention across time periods, and on first visit and longest visit to compare adherence of open access users with adherence of trial participants. During open access dissemination, most visits were first visits so that analyzing by longest visit yielded almost identical results.

Attrition describes the phenomenon that participants stop using the intervention [7]. Nonusage attrition refers to participants not returning to the intervention for repeated visits [7]. We also report attrition in terms of the duration of a single visit. In this sense, attrition refers to users who discontinue their visit at a specific point in time versus those who continue their visit beyond that point.

Comparisons are reported between trial participants (all of whom were registered, according to the study design), open access users (including both registered and unregistered open access users), and the subgroup of open access users who had registered and received a password to revisit Active-online.

Statistical Analyses

Demographic variables were compared between open access users and trial participants using *t*-tests and chi-square tests. Continuous variables measuring the use of Active-online (number of pages viewed and time spent in the tailored modules) were positively skewed. For these variables, therefore, the median and the interquartile range (IQR) are reported. The Wilcoxon-Mann-Whitney test and the chi-square test were used to compare use of Active-online between trial participants and open access users. Assuming that differences between time periods followed a time trend rather than a random pattern, a nonparametric test for trend across ordered groups developed by Cuzick [27] was performed for continuous variables, and a chi-square test for trend was performed for categorical variables to assess potential time trends across the six time periods. Logistic regression was used to assess potential predictors of repeated use of the website.

Nonusage attrition curves were based on the proportion of visitors still using the website up to a specific number of weeks or months after the first visit versus those who had stopped using it. The date of each user's last visit was designated as the date when program usage ended. The nonusage attrition curves are presented over 18 months (considered a suitable timeframe for trial participation) and over 12 weeks (for comparison with other published attrition curves). Similarly, attrition curves based on the duration of single visits (first visits and longest visits) are presented, which correspond to the proportion of visitors who had continued to use the intervention within a single session versus those who had ended the session after a specific number of minutes. Duration was defined as the time spent in the tailored modules as recorded in the user database;

therefore, individuals that did not enter a tailored module have been assigned a duration of zero. STATA 9.2 (STATACorp LP, College Station, TX, USA) was used for the analyses.

Results

User Characteristics

The yearly number of open access visits recorded in the Active-online database decreased from 42,626 in 2003-2004 to 8343 in 2008-2009. In the open access context, the proportion of women using Active-online increased from 49.5% in 2003-2004 to 61.3% in 2008-2009 (*P* [for trend] < .001). The mean age of open access users was between 38.4 years (95% confidence interval [CI] 38.0-38.8) in 2008-2009 and 40.4 years (95% CI 40.1-40.7) in 2005-2006 and 2006-2007. The proportion of open access users meeting the HEPA recommendations was between 39.9% in 2004-2005 and 42.6% in 2005-2006 (*P* [for trend] = .015).

Among open access Active-online users, 55.1% were women, while 74.9% of trial participants were women (*P* < .001). The mean age of open access users was 39.1 years (95% CI 39.0-39.2) compared with a mean age of trial participants of 43.1 years (95% CI 42.2-44.0) (*P* < .001). The proportion of individuals meeting the current Swiss HEPA recommendations did not differ significantly between open access users (40.9%) and trial participants (44.5%, *P*= .27).

Adherence to and Use of the Intervention

There were differences, but no consistent trends over time, in adherence to Active-online among open access users (based on the analysis of first visits). In general, use of the intervention among open access users was higher in 2003-2004, 2005-2006, and in 2006-2007, but lower in 2004-2005 and after 2007 (Table 2). Between 2003 and 2009, open access users who started a tailored module spent an average of 7.5 minutes in the program, with a median duration of 4.2 minutes. The subgroup of registered open access users who started a tailored module spent an average of 17.7 minutes in the program, with a median duration of 15.0 minutes.

For first visits, adherence to Active-online was highest for registered open access users (Table 2). Compared with all open access users, adherence was lower among trial participants. Trial participants visited fewer pages, and the proportion that started a tailored module, that spent at least 3 minutes in the modules, and that received at least one tailored feedback (HEPA or strength module) was smaller. However, trial participants who started a module tended to stay in the intervention longer.

Analyzing adherence according to longest visit, we found that results remained very similar to the results for first visit among open access users, indicating that the first and longest visit were identical among these users. This was not true for trial participants, however. As was the case for first visits, registered open access users achieved the highest adherence when results were based on longest visit. Trial participants' adherence was considerably higher during the longest visit than during the first visit (indicating that the first visit was not the longest visit) and was higher compared with all open access users.

Table 2. Adherence at first visit according to time periods during open access use 2003-2009, and at first and longest visit according to open access context and randomized controlled trial

	Median (IQR) Number of Pages Viewed per Visit	Number of Visits (%) When a Tailored Module Was Started	Number of Visits (%) When ≥ 3 Minutes Spent in the Tailored Modules	Number of Visits (%) When At Least One Tailored Feedback Was Received	Median (IQR) Minutes Spent in Tailored Module per Visit (When a Module Was Started)	Median (IQR) Minutes Spent in Tailored Modules per Visit (if ≥ 3 Minutes Spent in Modules)
Comparison of time periods[a] during open access use (based on first visits)						
2003-2004	16 (9-28)	29,967 (71.9%)	19,349 (46.4%)	24,973 (59.9%)	4.2 (1.8-11.4)	8.4 (4.8-15.6)
2004-2005	11 (4-21)	16,465 (65.8%)	9341 (37.3%)	13,132 (52.5%)	3.6 (1.2-8.4)	7.2 (4.2-13.8)
2005-2006	19 (11-31)	8851 (70.7%)	5593 (44.7%)	7277 (58.1%)	4.2 (1.8-11.4)	9.0 (4.8-15.6)
2006-2007	21 (13-32)	6661 (69.8%)	4154 (43.5%)	5716 (59.9%)	4.2 (1.8-10.8)	8.4 (4.8-15.6)
2007-2008	19 (11-30)	6015 (63.6%)	3728 (39.4%)	5152 (54.5%)	4.2 (1.8-10.2)	8.4 (4.8-14.4)
2008-2009	16 (11-25)	4818 (59.2%)	2639 (32.4%)	4090 (50.3%)	3.0 (1.2-7.8)	6.6 (4.2-11.4)
P Value[b]	< .001	< .001	< .001	< .001	< .001	< .001
Comparison of open access users and RCT participants (based on first visits)						
Open access: all users	16 (9-27)	72,777 (68.4%)	44,804 (42.1%)	60,340 (56.7%)	4.2 (1.8-10.2)	8.4 (4.8-15)
Open access: registered users only	42 (30-57)	4892 (96.2%)	4643 (91.3%)	4629 (91.1%)	15.0 (8.4-24.0)	15.6 (9.6-24.6)
Trial participants	7 (2-21)	322 (38.5%)	265 (31.7%)	250 (29.9%)	9.0 (3.6-15.6)	10.8 (6.0-16.2)
P Value[c]	< .001	< .001	< .001	< .001	< .001	< .001
Comparison of open access users and RCT participants (based on longest visits)						
Open access: all users	16 (9-27)	72,943 (68.5%)	44,985 (42.2%)	60,531 (56.8%)	4.2 (1.8-10.2)	8.4 (4.8-15.0)
Open access: registered users only	43 (31-57)	5021 (96.5%)	4789 (92.0%)	4779 (91.8%)	15.6 (9.0-24.6)	16.2 (10.2-25.2)
Trial participants	23 (8-38)	626 (74.4%)	554 (65.8%)	549 (65.2%)	12.0 (5.4-18.6)	13.2 (7.8-20.4)
P value[c]	< .001	< .001	< .001	< .001	< .001	< .001

[a]Each one-year time period started on April 23 and ended on April 22.

[b]*P* values are based onchi-square test for trend (categorical variables over time) and test for trend developed by Cuzick (continuous variables over time).

[c]*P* values for both the comparison between open access users (all) and trial participants, as well as between registered open access users and trial participants. *P* values are based on chi-square tests (comparisons between open access users and trial participants for categorical variables) and Wilcoxon-Mann-Whitney test (comparisons between open access users and trial participants for continuous variables).

Attrition

Figure 4 and Figure 5 show nonusage attrition curves for open access users (all open access users and subgroup of registered open access users only) and for trial participants over 18 months and over 12 weeks (for comparison with other published attrition curves, see [7,14]). The median lifetime website usage (time when 50% of users had stopped using the intervention) [13] was 0 days for open access users (all open access users and subgroup of registered open access users only) and 290 days for trial participants. In trial participants, the first two reminder emails after 9 and 10 months resulted in a relatively high proportion of individuals returning to Active-online by clicking on the link in the reminder email; however, fewer individuals returned after the third reminder. Reminder emails sent to registered open access users did not show the same effect: fewer than 6% were still using the website after the first reminder was sent out at two months.

Figure 4. Nonusage attrition curves for open access users and trial participants over 18 months

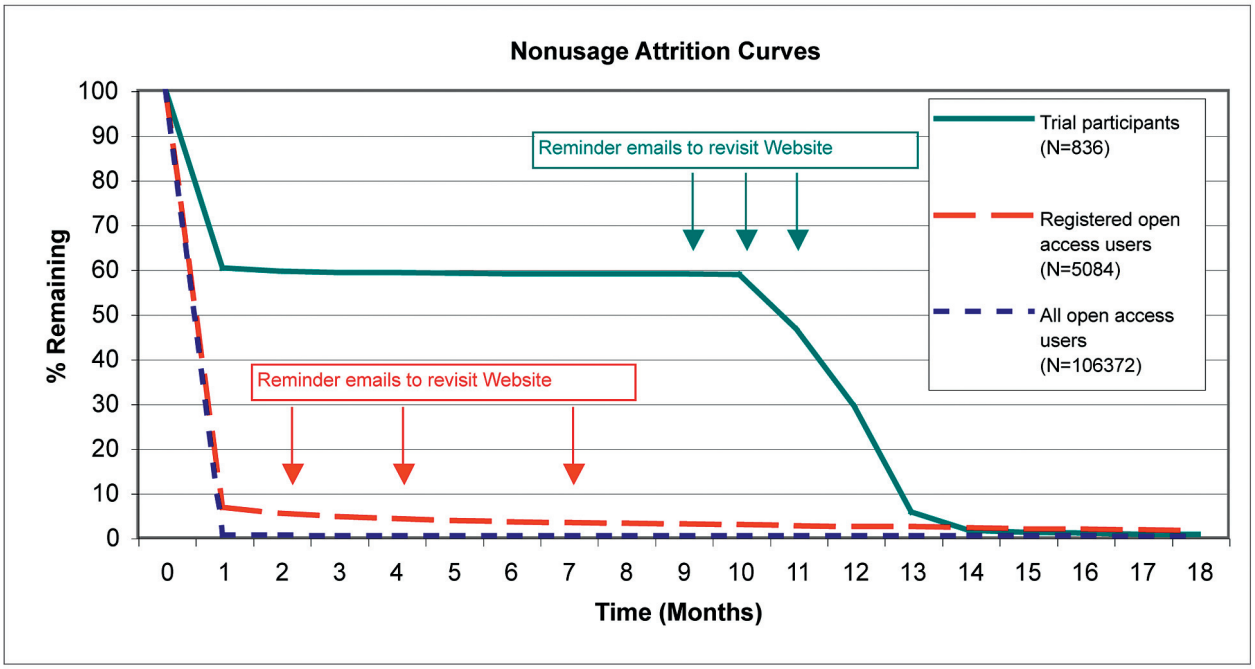


Figure 5. Nonusage attrition curves for open access users and trial participants over 12 weeks

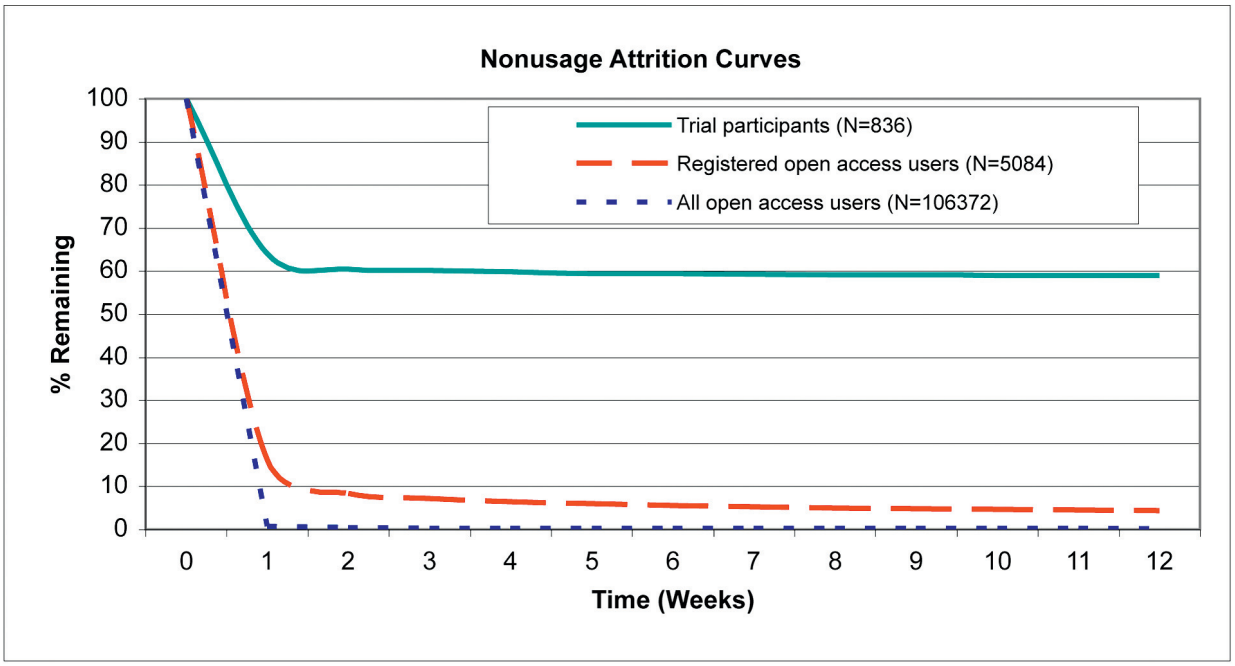


Figure 6 and Figure 7 present attrition curves for the time spent in the tailored modules during the first visit and during the longest visit, respectively, for open access users (all open access users and subgroup of registered open access users only) and for trial participants. Registered open access users spent more time in the tailored modules both during their first visit and during their longest visit compared with all open access users and compared with trial participants. The majority of trial participants did not spend much time in the tailored modules during their first visit (Figure 6); however, the proportion spending more time in the tailored intervention was much higher for longest visits (Figure 7). In contrast, the curves are very similar for first and longest visits of open access users (all open access users and subgroup of registered open access users only), indicating that first and longest visits were identical.

Figure 6. Attrition curves for the duration of the first visit for open access users and trial participants

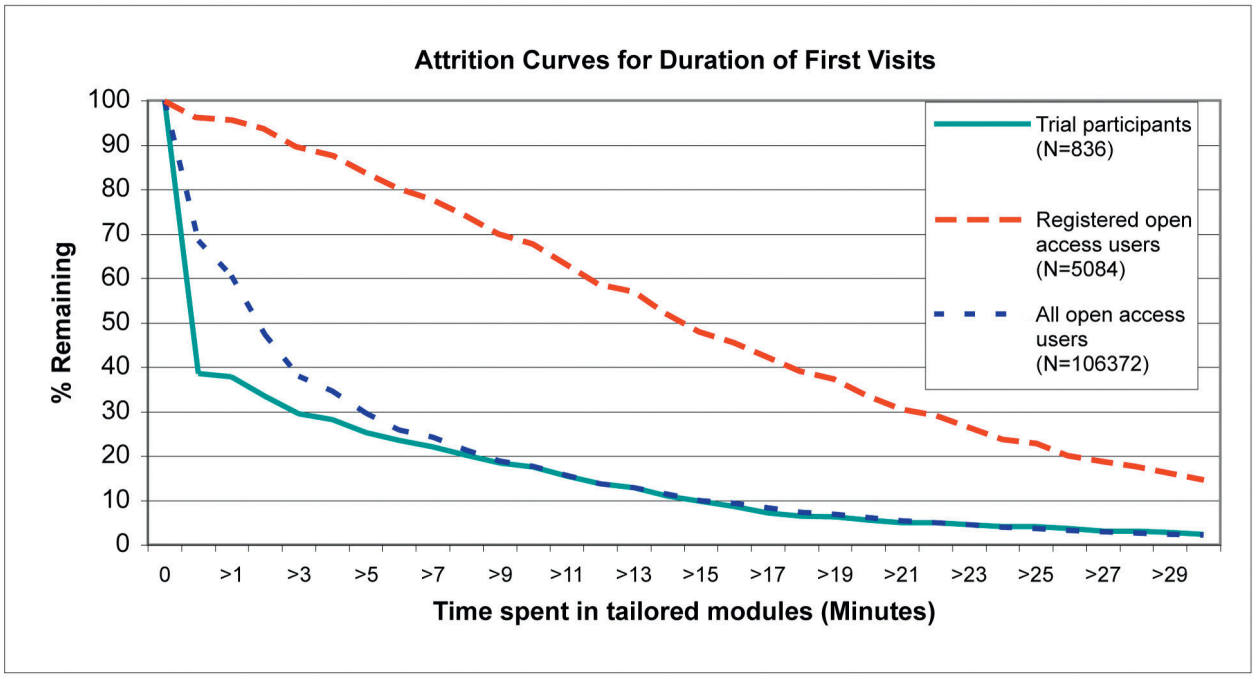
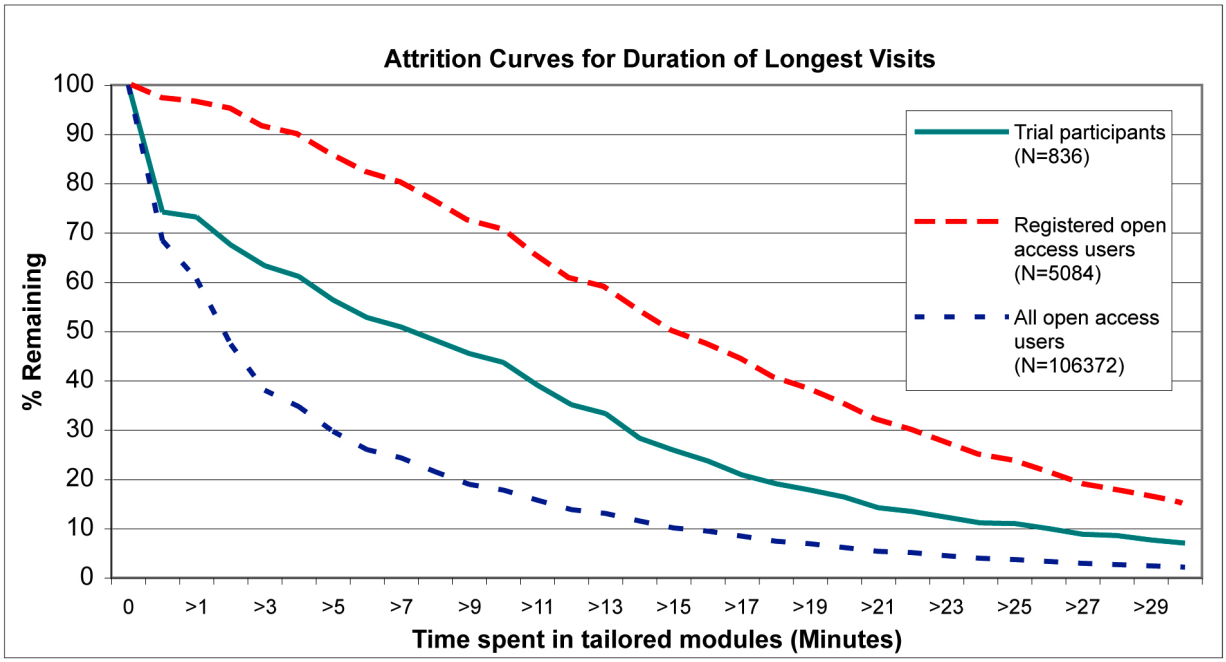


Figure 7. Attrition curves for the duration of the longest visit for open access users and trial participants



Predictors of Repeated Participation

In total, 1312 (25.8%) of open access users who registered and received a password and 558 (67.3%) of the trial participants returned for a repeated visit ($P < .001$). Table 3 displays potential predictors of repeated participation for registered open access users and for trial participants. Men and older individuals were significantly more likely to visit Active-online repeatedly in the open access context. Among trial participants, only the age group of 46-60 years (compared with < 30 years) was a significant predictor in the adjusted model, while gender did not predict repeated participation. Meeting the HEPA recommendations was not associated with repeated participation in registered open access users. In trial participants, however, there was a nonsignificant tendency for individuals not meeting the HEPA recommendations not to have returned for a repeated visit. Not having Swiss nationality achieved borderline significance as a predictor of lower rates of repeated participation in trial participants. Furthermore, there was a tendency for more highly educated individuals to have returned

for a repeated visit; however these associations were not significant. There were no effects for smoking and BMI.

Table 3. Predictors of repeated participation for registered open access users and for trial participants

	Registered Open Access Users (2003-2009)				Trial Participants			
	N	% Repeated Visits	Unadjusted OR (95% CI)	Adjusted OR[a](95% CI)	N	% Repeated Visits	Unadjusted OR (95% CI)	Adjusted OR[a](95% CI)
Gender								
Female	2,886	24.1	1.00	1.00	626	66.3	1.00	1.00
Male	2,197	28.0	*1.23* (1.08-1.39)	*1.20* (1.04-1.38)	210	70.5	1.21 (0.86-1.70)	0.79 (0.39-1.62)
Age category (years)								
< 30	1,270	20.5	1.00	1.00	151	53.6	1.00	1.00
30-45	2,111	25.9	*1.36* (1.15-1.60)	*1.37* (1.13-1.64)	324	65.7	*1.66* (1.12-2.46)	1.61 (0.71-3.66)
46-60	1,370	28.2	*1.52* (1.27-1.82)	*1.48* (1.21-1.81)	269	75.1	*2.61* (1.71-3.98)	*3.04* (1.25-7.38)
> 60	332	35.8	*2.17* (1.67-2.82)	*2.26* (1.68-3.04)	92	72.8	*2.32* (1.32-4.05)	1.72 (0.57-5.20)
Met HEPA recommendations								
Yes	1,587	24.4	1.00	1.00	347	70.3	1.00	1.00
No	2,697	25.8	1.07 (0.93-1.24)	1.11 (0.96-1.29)	489	65.2	0.79 (0.59-1.07)	0.76 (0.56-1.03)
Smoking	Not available							
Yes					125	64.8	1.00	1.00
No					711	67.8	1.14 (0.77-1.70)	1.03 (0.69-1.56)
BMI	Not available							
<= 25					501	67.5	1.00	1.00
25-30					227	68.7	1.06 (0.76-1.48)	0.93 (0.65-1.32)
> 30					106	64.2	0.86 (0.56-1.34)	0.79 (0.50-1.25)
Education	Not available							
Compulsory school					29	51.7	1.00	1.00
Apprenticeship					289	67.8	1.97 (0.91-4.24)	1.73 (0.78-3.81)
High school					124	63.7	1.64 (0.73-3.70)	1.82 (0.79-4.19)
Higher professional education, upper vocational school					189	70.9	*2.27* (1.03-5.03)	1.95 (0.86-4.41)
University					205	67.8	1.97 (0.90-4.31)	1.74 (0.78-3.91)
Nationality	Not available							
Swiss					737	68.4	1.00	1.00
Non Swiss					99	59.6	0.68 (0.44-1.05)	*0.64* (0.41-1.00)

[a]Adjusted for sex, age category, and whether HEPA recommendations were met. Additional adjustment for the other potential predictors in the model (RCT only) did not change the results.

Discussion

Principle Results and Comparison with Prior Work

The present study aimed to assess user characteristics, adherence, attrition, and predictors of repeated use in trial participants and open access users of a Web-based physical activity intervention. The most important findings were differences in adherence, attrition, and repeated participation between trial participants and open access users. Furthermore, reminder emails had a differential effect on attrition in trial participants and open access users. Assessing the data over time, there was an increase in the proportion of women using Active-online but no consistent trends in terms of adherence.

The yearly number of recorded visits on Active-online decreased from over 40,000 in 2003-2004 to less than 9,000 in 2008-2009. The most likely reason for this decrease was a decline in promotional efforts because there has been no active promotion of the website since 2008. Despite the decrease in the absolute number of visits, it is encouraging that even without active promotional strategies, Active-online still yielded around 23 visits per day in 2008-2009. Furthermore, no consistent time

trends in adherence and patterns of individual intervention use were observed in open access users between 2003 and 2009.

Different reasons may be responsible for the increase in the proportion of women using Active-online between 2003 and 2009. For one thing, the proportion of women using the Internet has increased steadily in Switzerland from 23% in 1997 to 44% in 2006 [28]. Moreover, women are generally more interested in health information and use a wider spectrum of information sources [29]. Specifically, "online" women are more likely to use the Internet to look for health information than "online" men [30,31].

Trial participants differed in several ways from open access users. Adherence of trial participants during the first visit was generally lower. Only the small proportion that became involved with the intervention spent as much or more time in the tailored modules compared with open access users. The additional baseline data assessment in the trial context is a likely reason for the low use during the first visit in trial participants. However, comparing trial participants and participants of the feasibility study [23], in the latter group adherence was higher and more similar to the patterns observed in open access users (data not shown). Therefore, adherence may vary in different controlled study settings.

Trial participants were significantly more likely to visit the website repeatedly compared with open access users. Furthermore, when analyzing longest visits (Table 2), we found that adherence was similar or higher in trial participants compared with all open access users. However, registered open access users still showed higher adherence during their longest visit. In a study comparing public registrants of a cognitive behavior therapy website with trial participants, the latter were more likely to adhere to the full treatment program [17,32]. In that study, trial participants were contacted weekly by phone, suggesting that the formal structure of the trial and the personal contacts may have increased compliance in trial participants compared with public registrants [17,32].

In the open access context, registration did not achieve high levels of repeated participation (Figure 4 and Figure 5). This indicates that reminder emails (with the same content) may not have the same effect in different contexts. Having agreed to participate in a study, trial participants may have felt more committed to react to reminder emails. We did find, however, that adherence during the first visit was significantly higher among registered open access users compared with unregistered users (Table 2). For example, there was a large difference in visit duration between registered open access users and all open access users. The differences in adherence were supported by the attrition curves for the duration of the first and longest visits (Figure 6 and Figure 7). The registration process itself is very brief and cannot explain the large difference in visit duration. Therefore, registered open access users seem to have been more motivated to use the intervention thoroughly compared with unregistered users.

Open access users and trial participants who started a tailored module spent an average of 7.5 and 9.2 minutes in the modules, respectively. Other studies have found similar visit durations, for example, an average of 9 minutes was found among participants of a randomized study regarding another physical activity website [9], an average of 7.1 minutes per visit was found on a tailored physical activity Internet intervention in a randomized study setting [11], and an average of 7 minutes was found for visitors of a smoking cessation website [33].

Only about 2% of the open access visits on Active-online between 2003 and 2009 were repeated visits (Table 1). A Web-based behavior change program for healthy body weight and healthy lifestyle, in which an email reminder strategy similar to ours was used, resulted in about 10% of users visiting the website more than once [20]. However, registration was compulsory for users, making it easier for the study investigators to detect repeated visits and possibly resulting in selection by more motivated users. Two smoking cessation websites yielded almost 20% [15] and 27% [16] of registered visitors returning to the website, respectively. When considering only registered users in our study, one quarter of the registered open access users and two thirds of the trial participants visited the intervention more than once.

In a previous study, the main predictors of repeated participation in a behavior change program for healthy body weight were older age, never having smoked, meeting the guidelines for moderate physical activity and vegetable consumption, and being obese [20]. In our study, older age was confirmed as a predictor of repeated participation, but smoking, BMI, and meeting HEPA recommendations were not. We did not obtain information about vegetable consumption. The only other significant predictors documented in our study were male sex among open access users (which was not significant in trial participants) and Swiss nationality, which achieved borderline significance in trial participants only. There was a tendency for more highly educated trial participants to have been more likely to return, an effect that was also reported for an online smoking cessation program [15]. In another study, repeated use of an interactive coaching program for smoking cessation was predicted by female sex and older age, among other smoking-related variables [16].

Nonusage attrition was much higher for open access users (all open access users as well as the subgroup of registered open access users only) than for trial participants. Similar, although less pronounced, results have been reported for spontaneous users of a cognitive behavior therapy website compared with participants in an RCT through the same website [17]. Attrition curves of open access visitors to other websites in the domains of cognitive behavior therapy [18] and the promotion of sensible drinking [19] were also comparable (see Ware et al [14], Figure 2, which shows different published attrition curves).

Limitations

The open structure of Active-online has advantages regarding dissemination and use of the intervention in that visitors are free to switch between modules, to open several windows concurrently, and to use the tailored intervention without registering. Thus, more individuals may be willing to participate in the intervention. However, this open structure also has some limitations. For example, a repeated visit of an unregistered user is recorded as a new first visit. Furthermore, an individual may open more than one tailored intervention browser window

resulting in multiple new visits being recorded in the database. Individuals are also free to stop the intervention at any point, which can produce large amounts of missing data if the intervention is terminated before all questionnaires are completed. Nevertheless, our study results provide insight into an open access Web-based physical activity intervention delivered under real-world conditions and allow comparisons of use and users over time and in different contexts.

Another limitation is the lack of information on the sociodemographic background of open access users and additional potential predictors of repeated website usage. During the development of Active-online it was decided not to include questions ascertaining sociodemographic variables (with the exception of sex and age) at the start of the tailored intervention. There was a concern that this may discourage entering the intervention for individuals who may be unwilling to reveal personal information or to spend time completing questions not related to tailored feedback. Ideally, a newer version of Active-online may include a brief questionnaire with questions related to smoking, BMI, socioeconomic status, education, and nationality, for example. Finally, we compared open access users visiting Active-online between 2003 and 2009 with trial participants visiting the website between 2006 and 2007. Thus, potential period effects may have influenced the differences observed in our analyses. However, Table 2 does not suggest specific time trends in adherence to the intervention; therefore, we think it was justifiable to use the full time range of data collected for open access users.

Conclusions

It is important to acknowledge that adherence, patterns of individual use, repeated participation, and attrition on a Web-based individually tailored physical activity intervention may differ between open access users and trial participants. Moreover, reminder emails to encourage repeated participation may not have the same effect in different contexts. These issues are important when interpreting and generalizing results of randomized controlled effectiveness trials.

Acknowledgments

Active-online was developed at the Institute of Social and Preventive Medicine at the University of Zurich and was funded and supported by the Swiss Federal Office of Sport, Health Promotion Switzerland, Allez Hop, Suva, and Qualitop. The authors would like to thank Thomas Suter who programmed the website.

Conflicts of Interest

None declared.

References

1. Oenema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. Ann Behav Med 2005;29(1):54-63. [Medline: 15677301]
2. Walters ST, Wright JA, Shegog R. A review of computer and Internet-based interventions for smoking behavior. Addict Behav 2006;31(2):264-277. [Medline: 15950392]
3. van den Berg MH, Schoones JW, Vliet Vlieland TPM. Internet-based physical activity interventions: a systematic review of the literature. J Med Internet Res 2007;9(3):e26 [FREE Full text] [Medline: 17942388] [doi: 10.2196/jmir.9.3.e26]
4. Vandelanotte C, Spathonis KM, Eakin EG, Owen N. Website-delivered physical activity interventions a review of the literature. Am J Prev Med 2007;33(1):54-64. [Medline: 17572313]
5. Oenema A, Brug J, Dijkstra A, de Weerdt I, de Vries H. Efficacy and use of an internet-delivered computer-tailored lifestyle intervention, targeting saturated fat intake, physical activity and smoking cessation: a randomized controlled trial. Ann Behav Med 2008;35(2):125-135. [Medline: 18363076]
6. 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] [Medline: 19403466] [doi: 10.2196/jmir.1194]
7. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [Medline: 15829473] [doi: 10.2196/jmir.7.1.e11]
8. 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] [Medline: 17513282] [doi: 10.2196/jmir.9.2.e11]
9. Leslie E, Marshall AL, Owen N, Bauman A. Engagement and retention of participants in a physical activity website. Prev Med 2005;40(1):54-59. [Medline: 15530581]
10. Sciamanna CN, Lewis B, Tate D, Napolitano MA, Fotheringham M, Marcus BH. User attitudes toward a physical activity promotion website. Prev Med 2002;35(6):612-615. [Medline: 12460529]
11. Lewis B, Williams D, Dunsiger S, Sciamanna C, Whiteley J, Napolitano M, et al. User attitudes towards physical activity websites in a randomized controlled trial. Prev Med 2008 Nov;47(5):508-513. [Medline: 18773915] [doi: 10.1016/j.ypmed.2008.07.020]

12. van den Berg MH, Ronday HK, Peeters AJ, Voogt-van der Harst EM, Munneke M, Breedveld FC, et al. Engagement and satisfaction with an Internet-based physical activity intervention in patients with rheumatoid arthritis. Rheumatology (Oxford) 2007 Mar;46(3):545-552 [FREE Full text] [Medline: 17043050] [doi: 10.1093/rheumatology/kel341]
13. 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] [Medline: 16954125] [doi: 10.2196/jmir.8.3.e15]
14. 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] [Medline: 19117828] [doi: 10.2196/jmir.1108]
15. Wang J, Etter JF. Administering an effective health intervention for smoking cessation online: the international users of Stop-Tabac. Prev Med 2004;39(5):962-968. [Medline: 15475030]
16. 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] [Medline: 19097975] [doi: 10.2196/jmir.1016]
17. 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] [Medline: 15631970] [doi: 10.2196/jmir.6.4.e46]
18. 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] [Medline: 15829479] [doi: 10.2196/jmir.7.1.e7]
19. 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] [Medline: 17513281] [doi: 10.2196/jmir.9.2.e10]
20. 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] [Medline: 17478410] [doi: 10.2196/jmir.9.1.e1]
21. Active-Online. URL: http://www.active-online.ch/
22. 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] [Medline: 19666456] [doi: 10.2196/jmir.1179]
23. Thüring N, Martin-Diener E, Martin B, Bauer G. Feasibility of a web-based randomised controlled trial for a tailored physical activity intervention. Schweiz Z Sportmed Sporttraumatol 2009;57(2):56-60 [FREE Full text]
24. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change. Applications to addictive behaviors. Am Psychol 1992;47(9):1102-1114. [Medline: 1329589]
25. Martin-Diener E, Thüring N, Melges T, Martin BW. The Stages of Change in three stage concepts and two modes of physical activity: a comparison of stage distributions and practical implications. Health Educ Res 2004 Aug;19(4):406-417 [FREE Full text] [Medline: 15155593] [doi: 10.1093/her/cyg051]
26. Martin BW, Mäder U, Stamm HP, Braun-Fahrländer C. Physical activity and health - what are the recommendations and where do we find the Swiss population? Schweiz Z Sportmed Sporttraumatol 2009;57(2):37-43 [FREE Full text]
27. Cuzick J. A Wilcoxon-type test for trend. Stat Med 1985;4(1):87-90 [FREE Full text] [Medline: 3992076] [doi: 10.1002/sim.4780040112]
28. Regelmässige Internetnutzung nach Geschlecht (Regular internet use according to gender). URL: http://www.bfs.admin.ch/bfs/portal/de/index/themen/21/02/ind32.indicator.70903.html [WebCite Cache ID 5cn8NFbKo]
29. Horch K, Wirz J. Nutzung von Gesundheitsinformationen (People's interest in health information). Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2005;48(11):1250-1255. [Medline: 16222408]
30. Fallows D. How Women and Men Use the Internet. Volume 2008. Washington, DC: Pew Internet & American Life Project; 2005. URL: http://www.pewinternet.org/~/media/Files/Reports/2005/PIP_Women_and_Men_online.pdf.pdf [WebCite Cache ID 5cqKUdopa]
31. 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] [Medline: 19275980] [doi: 10.2196/jmir.1035]
32. 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] [Medline: 14742346] [doi: 10.1136/bmj.37945.566632.EE]
33. Etter JF. Comparing the efficacy of two Internet-based, computer-tailored smoking cessation programs: a randomized trial. J Med Internet Res 2005;7(1):e2 [FREE Full text] [Medline: 15829474] [doi: 10.2196/jmir.7.1.e2]

Abbreviations

BMI: body mass index
CI: confidence interval
HEPA: health-enhancing physical activity

IQR: interquartile range
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 29.09.09; peer-reviewed by H Christensen, G Jerome; comments to author 22.10.09; revised version received 28.10.09; accepted 01.12.09; published 10.02.10.

Please cite as:
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
URL: http://www.jmir.org/2010/1/e3/
doi:10.2196/jmir.1361
PMID:20147006

Original Paper

Learning in a Virtual World: Experience With Using Second Life for Medical Education

John Wiecha[1], MD, MPH; Robin Heyden[2], BS; Elliot Sternthal[3], MD, FACP; Mario Merialdi[4], MD, PhD

[1]Department of Family Medicine, Boston University School of Medicine, Boston Medical Center, Boston, MA, USA

[2]Education Consultant, Wellesley, MA, USA

[3]Section of Endocrinology, Diabetes, and Nutrition, Department of Medicine, Boston Medical Center, Boston MA, USA

[4]Department of Reproductive Health and Research, World Health Organization, Geneva, Switzerland

Corresponding Author:
John Wiecha, MD, MPH
Department of Family Medicine
Boston University School of Medicine
Boston Medical Center
Dowling 5 South
1 BMC Place
Boston, MA 02118
Phone: 617 414 4465
Fax: 617 414 3345
Email: john.wiecha@bmc.org

Abstract

Background: Virtual worlds are rapidly becoming part of the educational technology landscape. Second Life (SL) is one of the best known of these environments. Although the potential of SL has been noted for health professions education, a search of the world's literature and of the World Wide Web revealed a limited number of formal applications of SL for this purpose and minimal evaluation of educational outcomes. Similarly, the use of virtual worlds for continuing health professional development appears to be largely unreported.

Methods: We designed and delivered a pilot postgraduate medical education program in the virtual world, Second Life. Our objectives were to: (1) explore the potential of a virtual world for delivering continuing medical education (CME) designed for physicians; (2) determine possible instructional designs using SL for CME; (3) understand the limitations of SL for CME; (4) understand the barriers, solutions, and costs associated with using SL, including required training; and (5) measure participant learning outcomes and feedback. We trained and enrolled 14 primary care physicians in an hour-long, highly interactive event in SL on the topic of type 2 diabetes. Participants completed surveys to measure change in confidence and performance on test cases to assess learning. The post survey also assessed participants' attitudes toward the virtual learning environment.

Results: Of the 14 participant physicians, 12 rated the course experience, 10 completed the pre and post confidence surveys, and 10 completed both the pre and post case studies. On a seven-point Likert scale (1, strongly disagree to 7, strongly agree), participants' mean reported confidence increased from pre to post SL event with respect to: selecting insulin for patients with type 2 diabetes (pre = 4.9 to post = 6.5, P= .002); initiating insulin (pre = 5.0 to post = 6.2, P= .02); and adjusting insulin dosing (pre = 5.2 to post = 6.2, P= .02). On test cases, the percent of participants providing a correct insulin initiation plan increased from 60% (6 of 10) pre to 90% (9 of 10) post (P= .2), and the percent of participants providing correct initiation of mealtime insulin increased from 40% (4 of 10) pre to 80% (8 of 10) post (P= .09). All participants (12 of 12) agreed that this experience in SL was an effective method of medical education, that the virtual world approach to CME was superior to other methods of online CME, that they would enroll in another such event in SL, and that they would recommend that their colleagues participate in an SL CME course. Only 17% (2 of 12) disagreed with the statement that this potential Second Life method of CME is superior to face-to-face CME.

Conclusions: The results of this pilot suggest that virtual worlds offer the potential of a new medical education pedagogy to enhance learning outcomes beyond that provided by more traditional online or face-to-face postgraduate professional development activities. Obvious potential exists for application of these methods at the medical school and residency levels as well.

(J Med Internet Res 2010;12(1):e1) doi:10.2196/jmir.1337

KEYWORDS

Medical education, continuing medical education, computer-assisted instruction, computer aided instruction, distance education, computer simulation, patient simulation, diabetes mellitus

Introduction

Background

Virtual worlds are rapidly becoming part of the educational technology landscape. Forterra's OLIVE, The Croquet Consortium, Sun Microsystem's Project Wonderland, ProtonMedia's Protosphere, and Linden Lab's Second Life are all examples of virtual world environments [1-5]. Platforms like these are potential environments for providing medical education.

Second Life (SL), one of the best known of these virtual worlds, consists of a flat-earth simulation of roughly 1.8 billion square meters, which would be about the size of Houston, Texas, if it were a physical place [6]. First launched in 2003, SL is an example of an immersive, three-dimensional environment that supports a high level of social networking and interaction with information. The SL virtual world is entered through a free client program called the SL viewer. Individuals enter SL as avatars that can take any form the user chooses. In the SL virtual world, residents can explore environments, meet and socialize with other residents (using voice and text chat), participate in individual and group activities, and learn from designed experiences. Built into the software is a three-dimensional modeling tool, based on simple geometric shapes that allows anyone to build virtual objects. These objects can be used, in combination with a scripting language, to add functionality [7].

While virtual worlds with their three-dimensional landscapes and customizable avatars seem similar to popular Massively Multiplayer Online Games (MMOGs), they do not adhere to the traditional definition of a game. Virtual worlds, like SL, are more focused on socializing, exploring, and building. As a result, there is an active educational community in SL. Over 300 colleges and universities have "builds" in SL where they teach courses and conduct research. A number of organizations, such as National Aeronautics and Space Administration (NASA), the National Oceanic and Atmospheric Association, National Institutes of Health, Jet Propulsion Laboratory, and National Public Radio, along with museums, educational groups, and a host of other government agencies, stage regular events, seminars, and workshops in Second Life [7].

Live sporting events, plays, meetings, seminars, research presentations, and musical concerts are all regular occurrences in Second Life. In 2008, Leong reported that the first virtual meeting of the International Virtual Association of Surgeons was held in Second Life, with 47 delegates attending from 5 countries [8]. There is no charge to create an SL account, although premium membership (which allows land ownership and greater technical support) is currently available for US$9.95/month. There is an economy in SL with the Linden dollar (L$) as the unit of exchange. Linden dollars can be used to buy, sell, rent or trade land, goods, or services.

Second Life demographics show that 83% of the population is 25 years or older, with the users over 44 years of age being the heaviest users on average. There is a close to even gender split among residents (57% male, 43% female), and more than 55% of the SL citizens come from outside the United States. In January 2008 residents spent a total of 28,274,505 hours in world and, on average, 38,000 residents were logged in at any particular moment. As of January 2009, just over 18 million accounts were registered (from over 150 countries), although there are no reliable figures for long-term, consistent usage [9,10,11,12,13].

Avatars are personalized by altering shapes, size, skin, hair, and clothing. They can be enlivened with animations to simulate facial expressions, posture, and gestures.

There are multiple channels for communication in the virtual world. Avatars can communicate by typing through local chat or instant messaging (IM), or by speaking through voice chat. Local chat is used for localized public conversation (called "local chat" or "backchat") between two or more avatars and is visible to any avatar within a given distance. IMs are used for private conversation and do not depend on local proximity.

The virtual world is not constrained by real-world physics. This is an important consideration when constructing educational activities. Avatars can fly, float to observe goings-on from any angle, teleport in order to materialize in a different location, fall and recover, and change their visual perspective at will.

Prior Work

Second Life has tremendous potential as a learning environment. The virtual world offers opportunities for student interaction, intense engagement, scripted immersive experiences, simulations, role-playing, and constructivist learning. The anonymity afforded by the avatar appears to lead to less inhibition and greater interaction. In addition, the greater sense of "presence" in a virtual world positively influences group process and cohesiveness, as well as engagement and attention [14,15,16]. We decided to use Second Life for our pilot since it is the most widely used virtual world platform and there is no charge to access it.

The problems with using a virtual world like SL for education and training lie mostly in the realm of technical and security issues. The software requires a download and has significant system requirements (processing power, up-to-date video card, and a fast broadband Internet connection) [17], the learning curve for navigation and interaction is steep, and the possibilities for technical problems and failures during the actual event are numerous. Many corporate or university firewalls do not allow access to public virtual worlds like Second Life.

Over the last three to four years, there has been growing interest in the medical and public health communities in using Second Life for public education, outreach, and training. There are a number of medical and health-related locations in Second Life. Most of these are education and awareness locations featuring

kiosks and visual displays, health videos, slideshows, and Web links [18-26]. Beard (2009) reports 68 relevant health-related SL locations, 34 of which were designed to disseminate health information [27]. There are several medical simulation sites where nursing or medical students can practice with virtual equipment, procedures, or lab results [28-32]. For example, the Imperial College of London has created a game-based simulation in Second Life for undergraduate medical students where they can interact with virtual respiratory therapy patients in order to build their skills and confidence [33,34]. Reports that describe the potential of Second Life for health professions education are common [18,35], but a recent comprehensive review [27] identified only 11 actual programs. Our search of English language peer-reviewed publication databases did not identify any formal evaluation of the educational effectiveness of health professional training in SL or other virtual worlds. Recent reports have commented on this lack of empirical evidence of learning impact [20,35]. We did find evaluated virtual world experiences in other disciplines, such as an interdisciplinary communications course taught in Second Life [36]. We concur with that article's author, Leslie Jarmon, that there are few empirical studies that inform instructional design and learning assessment in virtual worlds.

Physicians in the United States and in many other countries are required to undergo a specific number of hours of continuing medical education each year as a condition for maintenance of specialty board certification, and in some cases licensure, hospital admitting privileges, and insurance plan participation. The popularity and scope of available online, Web-based CME programs has increased dramatically in recent years [25,37,38], but has yet to include virtual worlds like SL as a venue.

Purpose

The purpose of this project was to explore the potential of using a virtual world platform for medical education through the development of a one-hour, interactive seminar for postgraduate primary care physicians on the topic of insulin therapy for type 2 diabetes.

The objectives of this pilot study were to: (1) explore the potential of the virtual world, Second Life, for CME activities; (2) determine possible instructional design approaches for using SL for CME; (3) understand the limitations of SL for CME activities; (4) understand the barriers, solutions, and costs to using SL, including participant and presenter training; and (5) measure participant learning outcomes and feedback.

The learning objectives for the participating physicians were to: (1) learn to be aware of insulin inertia (in other words, not delaying initiation of insulin therapy when indicated); (2) learn to use glycemic patterns to choose starting insulin; (3) learn to decide among basal, prandial, and premixed insulins; and (4) gain confidence titrating insulin.

Methods

The Venue

We started with an existing Boston University School of Medicine Second Life build (or "sim") [39] constructed on a private island owned and developed by Boston University School of Medicine (BUSM) for an earlier, joint project between BUSM and the World Health Organization (WHO). Ownership of the island allowed the developers to control access and thus provide security and privacy for the attending physicians. If an SL venue is not private, there is a risk of random avatars wandering into and potentially disrupting an event.

The existing BUSM/WHO meeting location was modified to add capacity, make it outdoors with no roof, open the walkways for easier navigation, and include automatic seating such that when avatars click on the seats, they automatically sit. We also built a media screen with a built-in script for controlling the PowerPoint slides that could also be used to project movies or websites.

Recruitment

Since this was a pilot, the plan was to keep the attendance small. Institutional Review Board approval was obtained from BUSM. Participants were recruited from two family medicine listservs and from prior participants in our online CME courses. Given the anticipated time commitment, including training and pre and post surveys, participants were offered an honorarium for completion of all activities. Since this was a pilot and not an accredited CME program, no CME credits were provided. Participants were offered a Second life coaching session if they had no previous SL experience.

Initially, 41 physicians expressed interest in the program. Ongoing email contact helped to further qualify the candidates, arrange coaching sessions, clarify objectives, confirm hardware/software requirements, and send candidates to the online pre survey. Over the course of three weeks, the original 41 was pared down to 14. Some had conflicts with the workshop date, some did not have a compatible computer, and still others did not respond to emails.

Of the 14 participants, 8 were female and all were primary care physicians (family medicine specialty). Participants had spent an average of 16 years in practice and resided in seven different states (NC, IL, CA, MA, SC, CT, KY). One observer attended the session logged in from Geneva, Switzerland. Of the 14 participants, 4 reported having previously logged into Second Life, and 9 described themselves as "daily Internet users."

The organizers commissioned Dr. Elliot Sternthal (MD, FACP, Director of Outpatient Diabetes Program, Boston Medical Center) to provide the session content. The instructional emphasis was to understand the options around diabetes patient assessment, insulin administration, and titrating dosage.

Training

Of the 14 participating physicians, 3 were experienced Second Life residents who did not require coaching. The remaining 11, plus the seminar speaker, were coached in one-on-one sessions conducted by phone with coach and participant each on their respective computers. It required an average of 12 email communications, and an average of 78 minutes of coaching time per participant to gain required proficiency. A checklist of skills (see Textbox 1) to be mastered was the organizing framework for each coaching session.

Textbox 1. Key Second Life skills taught in training sessions

Getting started

Logging on

Accessing the Second Life URL.

Do you have a computer gaming background?

Quit

Orientation

Camera controls: Understanding the various navigational commands allows residents to alter their viewpoint and focus in on specific items (eg, the slides in a presentation).

Mini map: This is a smaller map of the localized area in which a resident's avatar is standing that helps find other nearby avatars or landmarks.

Big map: This is a map of the entire Second Life grid. Residents can use this to locate themselves, to search, to teleport to new locations, and generate URLS that guide others to specific locations.

Create landmark: Sets a physical "bookmark" so that residents can return to the same location.

Environmental settings: Used to control the time of day.

Movement

Teleport: Used to transport from one location to another.

Walk

Fly

Right click on an object: Used to obtain more information about an object or another avatar.

Sit

Communicate

Audio preferences

Text chat

IM (Instant Message) a specific person.

Talk

Right click to see profile.

Offer friendship: Used to establish a connection with another avatar so that residents can easily communicate with them and exchange goods or services.

Use of Skype/ headset use.

Extras

Take a snapshot: Takes a photograph of whatever is currently on the screen.

Inventory management: This is the place where residents' virtual goods are stored (landmarks, clothing, objects, information cards). As the inventory grows, it is important to devise a filing system and a method for organizing it so that residents can easily find and retrieve desired items.

At the conclusion of the coaching session, participants were awarded a rating by one of the authors to indicate their degree of comfort with SL (1 = not comfortable, 2 = reasonably comfortable, 3 = very comfortable). The rating allowed us to partner confident doctors with less confident doctors. Partners were encouraged to arrange another meeting in SL with their partner prior to the workshop to practice navigational skills and camera controls. It was also hoped that knowing at least one other avatar would provide participants with a more comfortable experience at the actual event.

Instructional Design

The organizers worked with Dr. Sternthal to create an instructional design for the hour-long session and build an avatar that resembled him in real life. Dr. Sternthal started with a PowerPoint deck to support a 40-minute insulin therapy talk. We found that this deck, which would work well for a face-to-face talk, was too long and did not allow opportunities for the interaction and activity afforded by the virtual world. The organizers worked with Dr. Sternthal to shorten the talk, focus in on the key concepts, add visual elements, increase the interactivity, and leverage the unique capabilities of Second Life. Strategic questions were inserted (answers to be given by the doctors in local chat) in order to surface misconceptions and points of confusion among the participants. We also decided to introduce two mock diabetes patients to the session in order to apply the session content to a real-world scenario. This entailed designing two age-appropriate, overweight avatars. During the actual session, Dr. Sternthal gave a lecture interspersed with active engagement junctures. At those points, Dr. Sternthal asked questions related to his lecture, interviewed the patients,

displayed their lab results, and requested input from participant doctors on treatment plans.

Since the organizers and the speaker all lived in the same geographical area, we opted to conduct the seminar in the same physical location (see Figure 1) while the participating physicians were on their own computers in their homes or offices across the United States or in Switzerland. This allowed easy communication between the organizers during the event and convenient technical support for the speaker. As insurance against sound problems (one of the more often encountered SL technical issues), Skype [40] IDs were collected from all participants in advance so that a Skype conference call could be placed between the speaker's location and any participants experiencing sound problems.

Figure 1. CME event team on the night of the event

The Backchat

One of the features of Second Life is the ability to conduct text-based local conversations that every avatar in the immediate vicinity can "hear" (read). For this pilot, we experimented with the use of backchat in a number of ways. First we used it as a vehicle to increase workshop interactivity. At predetermined points throughout the seminar, Dr. Sternthal asked for the doctor's input (eg, "What would you recommend for this patient, given these lab values?") We also used backchat as a way to solicit more information from the participants (eg, "How many diabetics do you see in your practice?" "What are the most often-cited patient reasons for not wanting to go on insulin?") An unexpected advantage of the backchat that we noticed was that doctors talked to doctors throughout the event. Realizing the potential value of the backchat, we performed an analysis of its content. The seminar's entire chat log (25 pages) was printed and then coded according to the following categories:

Chatter: greetings, compliments, farewells, and casual conversation not pertaining to the content (eg, "Hi! Good to see you.")

Technical: questions or comments related to the functioning of Second Life (eg, "Remind me how to sit, please.")

Logistical: questions or statements related to the logistics of the event (pre and post surveys, timing, location, etc) (eg, "Yes, I did complete the survey.")

Doctors answering questions: statements given in answer to a question posed by the speaker (eg, "It depends on when the sugars are high.")

Doctors asking questions: questions posed to the speaker (eg, "When we start insulin, how often do we follow-up?")

Doctors exchanging information: instances of participants answering other participants' questions, providing links, making suggestions to each other (eg, "Is there a difference in A1C

between premixed insulin and NPH and regular?" Three other doctors answer, "Yes. Yes. Yes.")

Evaluation

Online surveys that included clinical skill confidence questions were completed before and following the session in order to assess physician reactions to the experience and evaluate learning transfer. These included two case studies before and two case studies after the session with questions designed to ascertain change in competence with respect to two learning objectives of the session: (1) how to properly initiate basal insulin in a patient with type 2 diabetes, and (2) how to initiate prandial (meal-time) insulin in such a patient. The cases were rated in a blinded fashion as either correct or incorrect by the two physician authors. This case approach has been shown to have good validity as compared to actual clinical practice [41].

Statistical Analyses

The significance of the increase in the proportion of correct scores on the case studies was tested using the Fisher exact test due to the small cell sizes. Fisher exact tests were performed using Epi Info version 3.5.1 (Centers for Disease Control and Prevention, Atlanta, GA) [42]. Means of responses on the Likert scale items were tested with paired *t* tests. The change from pre to post in the distribution of the Likert scale responses to the confidence questions was tested using the Wilcoxon signed rank test with the Statistix statistical package (Analytical Software, Tallahassee, FL) [43].

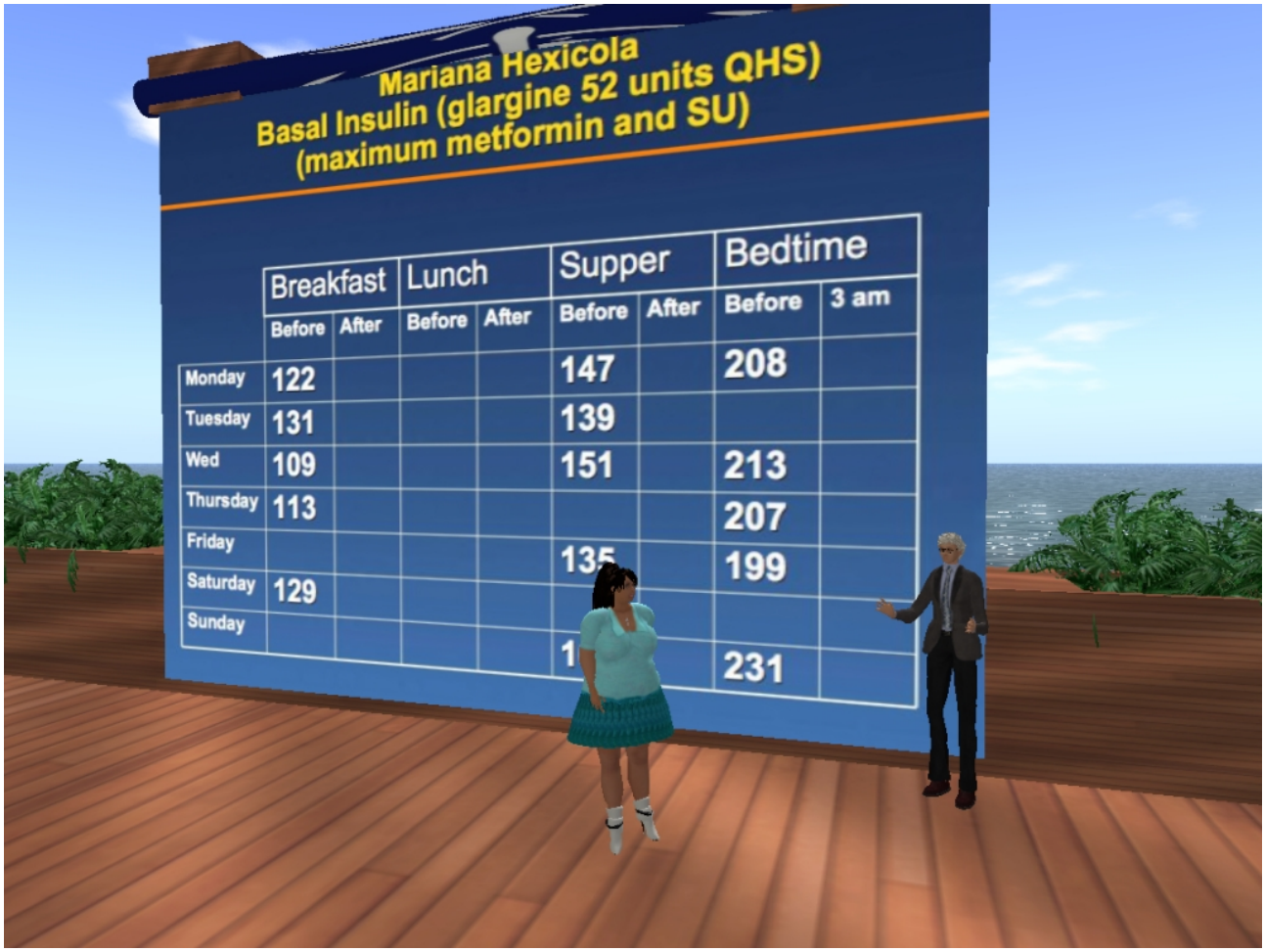
Results

The Event

The event was held on Monday, June 15, 2009, from 7:00pm to 8:00pm (EST). Avatars were asked to arrive 30 minutes early to get everyone settled and to resolve any technical issues. The session officially ended at 8:00pm, but the physicians stayed until 8:30pm asking questions and socializing. See Figures 2 and 3 for scenes from the actual event.

Figure 2. Avatars assembling just before the event on the Boston University/World Health Organization Second Life location

Figure 3. Dr. Elliot Sternthal, diabetes specialist, reviews lab results with fictional diabetic patient, Mariana Hexicola



Evaluation Outcomes

Of the 14 participant physicians, 12 rated the course experience, 10 completed the pre and post confidence surveys, and 10 completed both the pre and post case studies.

The data showing pre-post changes in confidence regarding insulin therapy can be found in Table 1. After participation in this program, the pilot group reported a statistically significant increase in confidence in their ability to select, initiate, and adjust insulin for patients with type 2 diabetes.

On test cases, the percent of participants providing a correct insulin initiation plan increased from 60% (6 of 10) pre to 90% (9 of 10) post (P= .2), and the percent of participants providing correct initiation of mealtime insulin increased from 40% (4 of 10) pre to 80% (8 of 10) post (P= .09).

All participants agreed that this SL experience was superior to other online methods, and most also felt that the SL method was as good as, if not better than, face-to-face methods. All agreed they would take other CME events in SL. Textbox 2 displays representative comments from the physicians about their experience with the seminar. Textbox 3 displays representative participant comments explaining why they felt the course was superior to face-to-face CME courses. Textbox 4 shares insights from Dr. Sternthal regarding his experience with the CME event.

Table 1. Course evaluation: change in confidence and competence and rating of course quality

Statement / Answer Categories	Pre Course	Post Course	*P* [a]
I am confident in my ability to select the appropriate insulin formulation and dose for starting a type 2 diabetic patient on insulin.			
Mean (N=10)	4.9	6.5	.002
Median	5	6.5	
Strongly agree	1	5	.02
Agree	2	5	
Agree somewhat	5	0	
Neutral	0	0	
Disagree somewhat	1	0	
Disagree	1	0	
Strongly disagree	0	0	
I am confident in my ability to initiate insulin therapy for my type 2 diabetes patients.			
Mean (N=10)	5.0	6.2	.02
Median	5	6	
Strongly agree	1	2	.04
Agree	3	8	
Agree somewhat	4	0	
Neutral	0	0	
Disagree somewhat	1	0	
Disagree	1	0	
Strongly disagree	0	0	
I am confident in my ability to appropriately adjust the insulin dose among patients with type 2 diabetes whom I have started on insulin.			
Mean (N=10)	5.2	6.2	.004
Median	6	6	
Strongly agree	1	4	.04
Agree	5	4	
Agree somewhat	1	2	
Neutral	1	0	
Disagree somewhat	2	0	
Disagree	0	0	
Strongly disagree	0	0	

Evaluation of Second Life Experience as Potential CME[b] Activity	Post Course	*P* [a]
Overall, I found this experience in Second Life to be an effective method of CME		
Strongly agree	6 (50%)	
Agree	4 (33%)	
Agree somewhat	2 (17%)	
Neutral	0	
Disagree somewhat	0	
Disagree	0	
Strongly disagree	0	
I would take other CME courses in Second Life.		

Evaluation of Second Life Experience as Potential CME[b] Activity	Post Course	*P* [a]
Strongly agree	9 (75%)	
Agree	3 (25%)	
Agree somewhat	0	
Neutral	0	
Disagree somewhat	0	
Disagree	0	
Strongly disagree	0	
The Second Life approach to CME was superior to other methods of online CME in which I have participated.		
Strongly agree	6 (50%)	
Agree	5 (42%)	
Agree somewhat	1 (8%)	
Neutral	0	
Disagree somewhat	0	
Disagree	0	
Strongly disagree	0	
This Second Life CME method is superior to face-to-face methods of CME.		
Strongly agree	2 (17%)	
Agree	1 (8%)	
Agree somewhat	3 (25%)	
Neutral	4 (33%)	
Disagree somewhat	2 (17%)	
Disagree	0	
Strongly disagree	0	
I would recommend to my colleagues to participate in a Second Life CME course.		
Strongly agree	6 (50%)	
Agree	6 (50%)	
Agree somewhat	0	
Neutral	0	
Disagree somewhat	0	
Disagree	0	
Strongly disagree	0	

[a]*P* value for change in means was determined with paired *t* test; *P* value for change in distribution of Likert scale responses was determined using Wilcoxon signed-rank test.

[b]This event was not a formally accredited continuing medical education (CME) activity.

Textbox 2. Qualitative feedback on the second life experience with medical education

”**What are the strengths of this potential method of CME?**

“Invites interaction without intimidation…”

“No travel, sitting at home. No cost…”

“Interactivity…”

“Felt like a more active participant…”

“Fun!”

“The local chat – the input from the other members made it more like a discussion group…”

“You only have to connect into one program and it is all there. It feels like a group experience.”

What were the most interesting elements of your experience?

“Loved sitting at home while attending live…”

“Real-time, ongoing questions, clinical comments.”

“The discussions with the ‘patients’…”

“The ability to see my colleagues questions in real-time ...and have them answered.”

“The anonymity...”

“The local chat/running comments offered a great way to ask questions, while thinking of them, without interrupting the speaker…”

What were the most confusing elements of your experience?

“Some of the local chat was distracting…”

“No confusing elements. Everything was great and in easy format.”

“Getting onto the SL site…”

“Getting going took some time. Once I had my avatar and basic directions, it was pretty easy.”

“At first, it was difficult to concentrate on the lecture while people were typing questions and comments.”

What are the weaknesses of this potential method of CME?

“Perhaps hard to convince other new users that the interface is easy — it really was — but there is a barrier to overcome.”

“None that I can think of.”

“New technology, still a bit raw…”

“Navigating in SL takes practice.”

“If internet connection doesn’t work…”

“Need participants who are tech savvy and able to multitask.”

“Definitely a new type of ‘manners’ and courtesy that needs to be learned in this type of CME.”

“Steep learning curve for some…”

What advice do you have for the developers?

“None.”

“Have friends available for chat/email.”

“Have some very good health sites to give us…”

“Perhaps not use local chat?”

“Some sound issues (echo)…”

“It was all fine. Lots of help when needed…”

“If not for the ‘guide at the side’ I would not have been able to participate.”

“Add pauses…time for questions, discourage ‘chatting.’” “Could have been a bit abbreviated.”

What did you learn about diabetes management that you didn’t know before?

“Adding bolus insulin to largest meal first, then reevaluating need to add bolus to other meals.”

“New perspective on newer agents.”

“I feel more comfortable with formulations of insulin now.”

Textbox 3. Ways in which this course was deemed superior to face-to-face CME events: participants' responses

"Able to listen to more views on the topic from other participants in a very time effective manner (by the typed answers for the presenters questions)."

"I would say that the biggest advantage is being able to attend the conference from wherever and still feel as if you were actually present at the conference."

"The presentation was very interactive. Instead of raising my hand and asking a question, I could type a question at any time and know that the presenter would eventually see it. I'm not the extroverted type, so don't always ask a question when given the opportunity."

"The 'patients' made it interesting and helped apply what we had learned to common patient scenarios. It's amazing how real it all feels during the presentation—a great way to learn."

"I feel having an alter ego on SL facilitates learning because of the freedom of expression—less fear of being judged. Unlike face-to-face meetings, looks, mannerisms, and speech patterns are taken out of the experience for the learner. Instead, the learner relaxes in a pleasant environment (home) and can concentrate on learning."

"This was one of the great experiences I have had. The phone or the other Web ones are a little boring. This one just kept me on my toes."

"There was a nice element of play that enhanced the fun aspect of learning. I miss having fun in class."

"Participants can communicate with each other during the session. No more saving questions till the end. The questions are addressed by all, not just the speaker, as they arise in the participants' minds. This enhances learning.

Textbox 4. Comments from Dr. Sternthal: experience conducting the event

This was indeed a unique experience for me. I have extensive experience in live face-to-face presentations to audiences of various sizes. I also have done numerous remote presentations via telephone and Web-based streaming where the audience is not visible just audible.

The SL presentation, while not face-to-face in the traditional sense, used a scenario that provided a sense of actually attending a meeting via the surrogate (avatar), thereby overcoming some of the sense of isolation (and often boredom) that is associated with just staring at remotely advanced slides on a monitor screen, typical of many remote programs. At the same time, with SL, there is a sense of anonymity and safety from critique and a loosening of inhibitions that facilitate the question and answer interaction. Therefore, SL seems to offer valuable features of both live face-to-face and remote programs: group interaction in a safe learning environment.

I would not consider moderating this SL program difficult. It was highly enjoyable and would not appear daunting to an experienced lecturer. I found certain challenges: coordinating the slides with the avatar patients, scanning and addressing in real-time the continuous chat questions and responses to clinical questions, and monitoring the overall flow of the program. It was important to keep things moving on time and to avoid prolonged silent times, so as to keep the participants engaged.

Interacting with the other avatars was not really different than fielding questions from a live audience. As I mentioned above, their presence gave a feeling of community and group dynamics to the presentation. The avatars all actively participated, likely because of the secure, non-judgmental learning environment.

The mock patients were invaluable and provided a concrete clinical correlate of the slide material. They reinforced the concepts being taught. In a way, they were like the live patients who used to grace medical lectures many years ago. However, since they are made up, there is no issue with confidentiality or privacy, and they can be constructed to portray the clinical situation that the lecture is addressing (in this program, deficiency in insulin initiation and titration).

I believe that learning was greatly facilitated by this SL virtual environment based upon the active participation, the number and quality of the questions, and the overall positive interactions with the presenter avatar and the other avatars.

The training that I received involved logistics in navigating to and through the SL virtual world, using and moving the avatar, and presenting information by speech, slides, and keyboard. This was accomplished in 3 sessions followed by a dry run of the whole presentation.

I believe that SL virtual learning is effective, innovative, stimulating, and fun. Its format encourages active reflection and participation, and the mock patients provide the opportunity to try out various clinical maneuvers, increasing the chance that the desirable behavioral change will be adopted by the health care participant. This has been a chronic problem with CME programs: effecting the appropriate behavioral change. It will be very valuable to evaluate the incorporation and durability of these take-home messages in participants after attending an SL medical education program.

Backchat Outcomes

In reviewing the backchat script, we recorded numerous incidences of doctors posing questions that were then answered by other participants, thus providing additional information over and above the seminar's planned content. One participant noted, "The input from the other members made it more like a discussion group…"

The results regarding the backchat are shown in Table 2, and examples of typical doctor-to-doctor exchanges are shown in Textbox 5. A significant amount of backchat was devoted to doctors answering questions from the speaker. This reflects the seminar's instructional design. In addition to this confirming information, there are a number of interesting trends revealed in the backchat. Almost all of the "chatter" (exchanges like "Hello, how are you?" "Thank you, it was great!") occurred at the beginning and end of the session. Almost all of the technical talk (eg, "How do I sit down?" "I can't hear.") happened in the first 30 minutes before the session started. The doctor-to-doctor exchanges were quite robust. For example, doctors provided insight from their own clinical practice, talked about specific patient examples, and pasted URLs into the backchat to help other doctors find further information online.

Table 2. Backchat communication during educational session

Topic of text communication during educational session	Number of lines posted by topic	% of total lines posted
Technical	221	32
Chatter	197	29
Doctors answering speaker	132	19
Logistics	65	9
Doctors asking speaker	52	8
Doctors exchanging	19	3
Total	686	100

Textbox 5. Representative backchat excerpt

Is the patient eligible for any assistance? Medicaid?

Maybe contact the drug companies; they have freebies for poor patients.

Could try premix NPH/regular.

Premixed 70/30 big, generic, cheaper than Lispro/Glargine combo.

Are there less expensive insulin formulas?

I agree with Abigail.

SU and metformin are $4/mon generic at certain pharmacies.

Any employer based options? A wellness program? Can we get the number of blood sticks down?

But she needs insulin.

While the backchat was used to increase the session interactivity, not all of the participants found it helpful. A few participants noted that it was distracting at times, and one suggested limiting the amount in future actual CME sessions. One possible explanation for the feeling of distraction is that, in SL, the backchat appears momentarily on top of the action on the screen. In future CME events, we should prepare the participants for this and provide advice for managing the information flow.

Developer Investment

Since one of the pilot objectives was to understand the time required to design and build a virtual world experience, careful record was kept of the organizers' hours. This is summarized in Table 3.

Table 3. Time spent developing and running the CME event

Item	Hours required for activity
Developer activities	
Planning	15
Instructional design	12
Coaching doctor participants	20
Venue build (work with builder, troubleshooting, creating signage)	10
Avatar design	10
Communications (email, telephone, meetings)	16
Technical issues/security	8
Rehearsals	6
Event	3
Project director effort	8
Speaker activities	
Rehearsals	6
Event	3
Total person-hours	109

Discussion

Principal Results

Overall, this pilot was very successful. The participant physicians' responses indicated that this was a positive and engaging experience that could meet their CME needs and fit their busy schedules. All respondents agreed it was superior to other methods of actual online CME they had experienced. In comparison with face-to-face CME, one-half of the respondents agreed Second Life CME was superior, one-third were neutral, and only two respondents disagreed "somewhat." This is a remarkable endorsement of such a new educational method. Our results showed that the virtual world model can have a positive impact on learner self-efficacy and, based on gains demonstrated from the cases, suggested a potential positive impact on clinical competence as well.

Many of the participants noted the convenience of an online seminar as one of the most important advantages—no travel is required, and they can participate from the comfort of their own homes. Since other online methods of instruction also offer this convenience, what justifies the expense of a virtual world course?

This pilot points to at least two important virtual world advantages: (1) the added sense of presence afforded by a representative avatar, and (2) the added real-life application provided by mock-patients [14,44]. Judging by the participant comments, the injected realism of mock patients was effective and added to the seminar's impact. These methods required considerable dexterity on the part of the speaker, Dr. Sternthal. The speaker must maintain fidelity to the planned script but, at the same time, pay attention to the backchat, adjust comments to answer participant questions, role play with the mock patient, and roll with whatever technical limitations, unexpected results, or problems arise. It will be worthwhile to explore and develop future instructional design options using mock patients.

Coaching Session Lessons

Relevant information emerged from the pre-event Second Life coaching sessions. There was tremendous variability in the participants' Second Life learning curve, with some quickly mastering the navigation and trying out functionality on their own, while others were very uncertain and hesitant to try buttons or commands. In hindsight, not all items on the skills checklist were needed for the event. Since the Second Life learning curve is so steep, it might be more effective to narrow the training to those skills absolutely required for the event and trust that the participant will go further with the virtual world on their own time, if they are interested.

Reasons for the SL orientation problems experienced by the doctors can be grouped into two categories, difficulty understanding the metaphor and technical skills:

1. Difficulty understanding the metaphor. Many doctors had trouble understanding where they were when training in the virtual environment. This was expressed by questions like, "Who is that?" (meaning the trainer's avatar), "What is it we're trying to do here?" "I'm confused about where I am," and repeated instances of participants going to the Second Life website [5] instead of opening the Second Life application in order to enter the virtual world.

2. Technical skills. For some doctors, mastering the menus, keyboard commands, and interface elements was no problem at all. Others had significant trouble remembering how to perform basic tasks such as sitting, chatting, accessing inventory, and setting landmarks. The variability in coaching time reflects these differences and is most likely related to the time each participant has spent online and/or working with other applications. It was noted that physicians with gaming experience were much more adept at mastering the basics.

It was also interesting to note the differences in approach to learning. Some of the physicians needed to understand the context for any navigational command ("Why would you need to take a snapshot in Second Life?" or "When would you use that function?"), while others were perfectly content to run through a list of skills and tick them off with no need to attach relevance.

We observed that subjects tended to choose avatars that looked like them in real life. For example, one African American participant explained, "But I need the avatar's skin to be darker.…" This is consistent with previous studies [44,45] showing that people choose avatars for self-representation, based primarily on how similar avatars are to themselves. It also implies the increased sense of presence afforded by the virtual world [14].

The tactic of partnering confident with less confident subjects for additional practice met with modest success but could be used more effectively in the future if there was more time between the coaching session and the event.

Overall Lessons Learned

Since this pilot was primarily designed to develop expertise in designing and running CME and other medical education events programs in the virtual world, it is important to capture key lessons for future programs.

Among the lessons learned is that an event like this has to be designed in such a way that it answers the question "Why SL?" before it gets asked. In other words, it is critical to take advantage of the unique opportunities afforded by the virtual world to push the experience beyond what could be delivered via a website or a webinar.

With so many opportunities to "go wrong" with a technically complex platform, it is clear that thorough planning is paramount. This event was meticulously planned, rehearsed, and buttressed with contingency plans.

The time investment to plan, prepare, and deliver a one-hour event is steep, and the learning curve for the doctor-participants is not trivial. From our experience, delivering a CME event in Second Life is much more time consuming than face-to-face or webinar events.

Virtual world events require skilled and unflappable speakers. Speakers must be knowledgeable, confident in their expertise, unruffled with the inevitable technical glitches, good humored,

in tune with the students, and able to field questions from the backchat while still keeping an eye on content and timing.

Overall, there is a certain excitement that comes from just being there, all together, in the virtual world. The physician participants are more forthcoming, brave, involved, and present. There is an interesting ”protection“ effect that comes with the avatar anonymity that promotes disclosure and sharing. Many of the doctors commented on it, and from the backchat logs, one could certainly sense an openness and willingness to venture further. In fact, the informal exchange and learning resulting from the backchat suggests a less top-down and more Socratic approach to learning and instruction with participants and experts exchanging strategies and ideas more fluidly. What's more, the backchat exchange might result in enhanced professional connections among participants that could extend beyond the boundaries of the single event.

But while the backchat offers an opportunity for increased interactivity, it can be distracting. Organizers should prepare participants for this and set ground rules about keeping the chat to a professional level.

Lastly, it is important to leverage the more playful aspects of Second Life without drifting into unprofessional behavior. In this event, the mock-patient avatars' responses, the 15-minute conversational warm-up before the session started, and a social phase (including “serving” champagne) at the end all helped to serve that purpose.

Comparison with Prior Work

Prior reports have described the potential of SL and other virtual worlds for the training of health professionals, or have described a curriculum or program in SL [24,33,34,46,47,48]. Most of these reports, however, have not included feedback from the health professionals who were involved in this training. One report described improvement in leadership skills after participation in a trauma exercise in a non-Web-based virtual environment created by the investigators for a local client [32].

However, we did not locate prior peer-reviewed publications that described the impact of training delivered in virtual worlds like SL on health professionals' confidence or on health professionals' knowledge and competence in clinical management.

Limitations

Chief among concerns for future CME programming is the time required to bring all participants to a functional level and the technical requirements/potential problems with the software. The time investment to scale the course must be considered. One hour per participant in one-on-one coaching sessions is not scaleable to larger enrollments. Future events will require group or self-directed training, and the training should be limited to the skills required for the event. While the technical requirements to run a virtual world are steep, it is clear that ever-increasing processing power and bandwidth will make this less of an issue going forward. Security concerns around virtual worlds will most likely be addressed with the development of platforms for installation behind firewalls and integration with other native systems. The potential for technical failure in this particular virtual world platform is high. System crashes and sound problems are the most commonly experienced issues. It is essential to have back-up plans for each possible technical snafu.

Although the improvements in performance on the case scores were large, our small sample size limited our ability to assess the statistical significance of these score changes. More robust evaluation work with larger samples is needed to measure impact on clinical skills and outcomes and to compare the relative effectiveness and efficiency of educational methods. Such work also needs to address the rapidly developing diversity of online approaches including virtual worlds. Also, further careful evaluation work is needed on less selected populations to determine if the enthusiastic reception of this SL method is replicable more broadly.

Conclusions

The organizers are very optimistic about future use of virtual worlds applied to all phases of medical education including continuing medical education. The enriched environment, convenience, and possibilities for constructivist approaches all add up to tremendous potential. What's more, it is expected that virtual worlds will become more ubiquitous in other types of computing. Some speculate that virtual worlds will soon replace our Internet browsers [30].

There are many instructional design possibilities to consider that will further leverage the unique advantages of these innovative technological environments [49]. In future events, we will consider expanding the mock patient element so that physician pairs (or teams) could interview their own mock patients and compare findings. This should enhance learning on the psychosocial aspects of patient care, allowing doctors to interact with other avatars (including patients, staff, and experts) in a safe, simulated environment and reflect on their learning. The backchat could be used more specifically to boost engagement and information transfer. Simulated physiology models could be created to provide a deeper understanding of the biology behind the disease state. We intend to further investigate the unique potential of Second Life as a learning environment.

The results of this pilot suggest that virtual worlds offer tremendous opportunity to provide a space for constructivist learning at its best and to enhance learning outcomes beyond that provided by traditionally designed CME courses.

Acknowledgments

This work was supported by a grant from The Physicians' Foundation to Dr. Wiecha.

References

1. Forterra Systems. OLIVE URL: http://www.forterrainc.com [accessed 2009 Sep 8] [WebCite Cache ID 5iseHBK8R]
2. Croquet Consortium. URL: http://www.opencroquet.org/index.php/Main_Page [accessed 2009 Sep 9] [WebCite Cache ID 5iseWKi8b]
3. Sun Microsystems. Project Wonderland URL: https://lg3d-wonderland.dev.java.net [accessed 2009 Sep 8] [WebCite Cache ID 5isedCRJb]
4. Proton Media. Protosphere URL: http://www.protonmedia.com [accessed 2009 Sep 8] [WebCite Cache ID 5k9aRnbo8]
5. Linden Labs. Second Life URL: http://secondlife.com [accessed 2009 Sep 8] [WebCite Cache ID 5iseoDiN5]
6. Linden Labs. Why work in Second Life? URL: http://work.secondlife.com/whyslwork/whatis/ [accessed : 2009 Sep 9] [WebCite Cache ID 5jpohATli]
7. EDUCAUSE Learning Initiative. Seven things you should know about Second Life URL: http://www.educause.edu/ELI/7ThingsYouShouldKnowAboutSecon/163004 [accessed 2009 Sep 8] [WebCite Cache ID 5isfVuCpk]
8. Leong JJ, Kinross J, Taylor D, Purkayastha S. Surgeons have held conferences in Second Life. Br Med J 2008;337:a683. [Medline: 18614495] [doi: 10.1136/bmj.a683]
9. Linden Labs. Second Life User Statistics URL: https://blogs.secondlife.com/community/features/blog/2009/01/15/second-life-residents-logged-nearly-400-million-hours-in-2008-growing-61-over-2007 [accessed 2009 Sep 13] [WebCite Cache ID 5k9acti1M]
10. Linden Labs. Second Life Usage Demographics URL: http://www.bananaverse.com/2008/05/07/second-life-usage-demographics [accessed 2009 Sep 8] [WebCite Cache ID 5k9aqWwlO]
11. Wagner JA. Gigaom. Exclusive: Second Life starts to grow again URL: http://gigaom.com/2009/04/15/exclusive-internal-second-life-data-shows-returning-growth [accessed 2009 Sep 8] [WebCite Cache ID 5isfPKIHI]
12. Johnson L, Levine A, Smith R. The 2009 Horizon Report. The New Media Consortium 2009 Jan [FREE Full text] [WebCite Cache ID 5jpoXya6K]
13. Cross J, O'Driscoll T, Trondsen E. Another life virtual worlds as tools for learning. eLearn Magazine 2007;22 [FREE Full text] [WebCite Cache ID 5jpp48YSa]
14. Zhao S. Toward a taxonomy of copresence. Presence: Teleoperators and Virtual Environments 2003;12(5):445-455. [doi: 10.1162/105474603322761261]
15. Yee N, Bailenson JN, Urbanek M, Chang F, Merget D. The unbearable likeness of being digital: the persistence of nonverbal social norms in online virtual environments. Cyberpsychol Behav 2007 Feb;10(1):115-121. [Medline: 17305457] [doi: 10.1089/cpb.2006.9984]
16. Taylor TL. Living digitally: embodiment in virtual worlds. In: The Social Life of Avatars: Presence and Interaction in Shared Virtual Environments. London, UK: Springer-Verlag; 2002. URL: http://tltaylor.com/wp-content/uploads/2009/07/Taylor-LivingDigitally.pdf
17. Linden Labs. Second Life system requirements URL: http://secondlife.com/support/sysreqs.php [accessed 2009 Sep 8] [WebCite Cache ID 5isfEaE9J]
18. Boulos MN, Hetherington L, Wheeler S. Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. Health Info Libr J 2007;24(4):233-245. [Medline: 18005298] [doi: 10.1111/j.1471-1842.2007.00733.x]
19. Gorini A, Gaggioli A, Vigna C, Riva G. A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology. J Med Internet Res 2008;10(3):e21 [FREE Full text] [Medline: 18678557] [doi: 10.2196/jmir.1029]
20. Hansen MM. Versatile, immersive, creative and dynamic virtual 3-D healthcare learning environments: a review of the literature. J Med Internet Res 2008;10(3):e26 [FREE Full text] [Medline: 18762473] [doi: 10.2196/jmir.1051]
21. Huang ST, Kamel Boulos MN, Dellavalle RP. Scientific discourse 2.0. Will your next poster session be in Second Life? EMBO Rep 2008 Jun;9(6):496-499 [FREE Full text] [Medline: 18516082] [doi: 10.1038/embor.2008.86]
22. Woodruff P. Medicine's not-so-secret Second Life: public health education thrives in so-called "virtual worlds. National Review of Medicine 2007;30(6):4.
23. ScienceRoll. Top ten virtual medical sites in Second Life URL: http://scienceroll.com/2007/06/17/top-10-virtual-medical-sites-in-second-life [accessed 2009 Sep 9] [WebCite Cache ID 5jplPuDdj]
24. Boulos MNK, Toth-Cohen S. The University of Plymouth sexual health SIM experience in Second Life: evaluation and reflections after one year. Health Info Libr J 2009 Jun 11;26(4):279-288. [Medline: 19930475]
25. iReport.com. 2008 Nov 19. AMIA Panel Discussion on virtual worlds URL: http://www.ireport.com/docs/DOC-149486 [WebCite Cache ID 5k9bPHN40]
26. Schmidt B, Stewart S. Implementing the virtual reality learning environment: Second Life. Nurse Educ 2009;34(4):152-155. [Medline: 19574850] [doi: 10.1097/NNE.0b013e3181aabbe8]
27. Beard L, Wilson K, Morra D, Keelan J. A survey of health-related activities on second life. J Med Internet Res 2009;11(2):e17 [FREE Full text] [Medline: 19632971] [doi: 10.2196/jmir.1192]
28. Webicina. Practicing medicine in the web 2.0 era URL: http://www.webicina.com/medicine_in_second_life/what_is_second_life_75 [accessed 2009 Sep 13] [WebCite Cache ID 5k9bfYXWR]

29. Kamel Boulos MN, Ramloll R, Jones R, Toth-Cohen S. Web 3D for public, environmental and occupational health: early examples from second life. Int J Environ Res Public Health 2008 Dec;5(4):290-317 [FREE Full text] [Medline: 19190358] [doi: 10.3390/ijerph5040290]
30. Toro-Troconis M, Boulos MNK. Musings on the state of 3D virtual worlds for health and healthcare. Journal of Virtual Worlds Research 2009;2(2) [FREE Full text] [WebCite Cache ID 5jplfu2SS]
31. Lafsky M. Discover Magazine. Can training in Second Life teach doctors to save real lives? URL: http://discovermagazine.com/2009/jul-aug/15-can-medical-students-learn-to-save-real-lives-in-second-life/article_view?b_start:int=1&-C= [accessed 2009 Sep 13] [WebCite Cache ID 5k9bplYmv]
32. Heinrichs LW, Youngblood P, Harter PM, Dev P. Simulation for team training and assessment: case studies of online training with virtual worlds. World J Surg 2008 Feb;32(2):161-170. [Medline: 18188640] [doi: 10.1007/s00268-007-9354-2]
33. Toro-Troconis M, Mellström U, Partridge M, Meeran K, Barrett M, Higham J. Designing game-based learning activities for virtual patients in Second Life. J Cyber Rehabil 2008;1(3):225-238 [FREE Full text] [WebCite Cache ID 5jppGcDCT]
34. Mesko B. ScienceRoll. 2009 Jun. Imperial College London in Second Life: Virtual Patients URL: http://scienceroll.com/2009/06/11/imperial-college-london-in-second-life-virtual-patients/ [WebCite Cache ID 5k9c6W6yT]
35. Hansen MM, Murray PJ, Erdley WS. The potential of 3-D virtual worlds in professional nursing education. Stud Health Technol Inform 2009;146:582-586. [Medline: 19592909]
36. Jarmon L, Traphagan T, Mayrath M, Trivedi A. Virtual world teaching, experiential learning, and assessment: an interdisciplinary communication course in Second Life. Comput Educ 2009 Aug;53(1):169-182. [Medline: 19632971]
37. Association of American Medical Colleges. 2008. Academic CME in North America: The 2008 AAMC/SACME Harrison Survey URL: http://www.sacme.org/site/sacme/assets/pdf/2008_CME_Report.pdf [WebCite Cache ID 5jpqED327]
38. Society for Continuing Medical Education. 2007 Jan. Society for Continuing Medical Education Survey for 2006 URL: http://www.sacme.org/site/sacme/assets/pdf/BS_results_2006.pdf [WebCite Cache ID 5jpqJaATn]
39. World Health Organization Boston University Medical School. Second Life location URL: http://slurl.com/secondlife/Teaching%209/12/169/22 [accessed 2009 Sep 8] [WebCite Cache ID 5lu04c5gO]
40. Skype. Skype: Internet phone software URL: http://skype.com [accessed 2009 Sep 12] [WebCite Cache ID 5lu0W0I13]
41. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. JAMA 2000 Apr 5;283(13):1715-1722 [FREE Full text] [Medline: 10755498] [doi: 10.1001/jama.283.13.1715]
42. Centers for Disease Control and Prevention. Epi Info Version 3.5.1 URL: http://www.cdc.gov/epiinfo/ [accessed 2009 Dec 02] [WebCite Cache ID 5ljMzFtcs]
43. Analytical Software. Statistix 9 URL: http://www.statistix.com/ [WebCite Cache ID 5mmMCiWJB]
44. Messinger P. On the relationship between my avatar and myself. Journal of Virtual Worlds Research 2008;1(2) [FREE Full text] [WebCite Cache ID 5k9cYL32a]
45. Nowak KL. The influence of the avatar on online perceptions of anthropomorphism, androgyny, credibility, homophily, and attraction. Journal of Computer-Mediated Communication 2005;11(1) [FREE Full text] [WebCite Cache ID 5lu0eyyy1]
46. Gerald S, Antonacci DM. Virtual world learning spaces: developing a Second Life operating room simulation. Educause Quarterly 2009;http.
47. Stewart S, Pope D, Duncan D. Using Second Life to enhance ACCEL an online accelerated nursing BSN program. Stud Health Technol Inform 2009;146:636-640. [Medline: 19592919]
48. Toro-Troconis M, Partridge M, Mellstrom U, Meera K, Higham J. Technical infrastructure and initial findings in the design and delivery of game-based learning for virtual patients in Second Life. In: Proceedings of Researching Learning in Virtual Environments International Conference. 2008 Presented at: Researching Learning in Virtual Environments International Conference; November 20-21, 2008 p. 334-350.
49. Cannon-Bowers JA. Recent advances in scenario-based training for medical education. Curr Opin Anaesthesiol 2008 Dec;21(6):784-789. [Medline: 18997530] [doi: 10.1097/ACO.0b013e3283184435]

Edited by G Eysenbach; submitted 14.08.09; peer-reviewed by L Beard, L Johnson, P Yellowlees; comments to author 02.09.09; revised version received 09.11.09; accepted 18.11.09; published 23.01.10.

Please cite as:
Wiecha J, Heyden R, Sternthal E, Merialdi M
Learning in a Virtual World: Experience With Using Second Life for Medical Education
J Med Internet Res 2010;12(1):e1
URL: http://www.jmir.org/2010/1/e1/
doi:10.2196/jmir.1337
PMID:20097652

Original Paper

Physician Order Entry Or Nurse Order Entry? Comparison of Two Implementation Strategies for a Computerized Order Entry System Aimed at Reducing Dosing Medication Errors

Alireza Kazemi[1,2,3], MD, PhD; Uno GH Fors[1], DDS, PhD; Shahram Tofighi[4], MD, PhD; Mesfin Tessma[1], MD, MPH; Johan Ellenius[1], PhD

[1]Department of Learning, Informatics, Management and Ethics (LIME), Karolinska Institutet, Stockholm, Sweden

[2]Management Information Systems Centre, Hamadan University of Medical Sciences, Hamadan, Iran

[3]National Public Health Management Centre (NPMC), Tabriz, Iran

[4]Research Centre for Strategic Studies on Health Care, Baqyatallah University of Medical Sciences, Tehran, Iran

Corresponding Author:
Alireza Kazemi, MD, PhD
Medical Informatics Group, Department of Learning, Informatics, Management and Ethics (LIME)
Karolinska Institutet
Berzelius väg 3
SE - 171 77, Stockholm
Sweden
Phone: +46 8 52483926
Fax: +46 8 52483600
Email: alireza.kazemi@ki.se

Abstract

Background: Despite the significant effect of computerized physician order entry (CPOE) in reducing nonintercepted medication errors among neonatal inpatients, only a minority of hospitals have successfully implemented such systems. Physicians' resistance and users' frustration seem to be two of the most important barriers. One solution might be to involve nurses in the order entry process to reduce physicians' data entry workload and resistance. However, the effect of this collaborative order entry method in reducing medication errors should be compared with a strictly physician order entry method.

Objective: To investigate whether a collaborative order entry method consisting of nurse order entry (NOE) followed by physician verification and countersignature is as effective as a strictly physician order entry (POE) method in reducing nonintercepted dose and frequency medication errors in the neonatal ward of an Iranian teaching hospital.

Methods: A four-month prospective study was designed with two equal periods. During the first period POE was used and during the second period NOE was used. In both methods, a warning appeared when the dose or frequency of the prescribed medication was incorrect that suggested the appropriate dosage to the physicians. Physicians' responses to the warnings were recorded in a database and subsequently analyzed. Relevant paper-based and electronic medical records were reviewed to increase credibility.

Results: Medication prescribing for 158 neonates was studied. The rate of nonintercepted medication errors during the NOE period was 40% lower than during the POE period (rate ratio 0.60; 95% confidence interval [CI] .50, .71;$P < .001$). During the POE period, 80% of nonintercepted errors occurred at the prescription stage, while during the NOE period, 60% of nonintercepted errors occurred in that stage. Prescription errors decreased from 10.3% during the POE period to 4.6% during the NOE period ($P < .001$), and the number of warnings with which physicians complied increased from 44% to 68% respectively ($P < .001$). Meanwhile, transcription errors showed a nonsignificant increase from the POE period to the NOE period. The median error per patient was reduced from 2 during the POE period to 0 during the NOE period ($P = .005$). Underdose and curtailed and prolonged interval errors were significantly reduced from the POE period to the NOE period. The rate of nonintercepted overdose errors remained constant between the two periods. However, the severity of overdose errors was lower in the NOE period ($P = .02$).

Conclusions: NOE can increase physicians' compliance with warnings and recommended dose and frequency and reduce nonintercepted medication dosing errors in the neonatal ward as effectively as POE or even better. In settings where there is major physician resistance to implementation of CPOE, and nurses are willing to participate in the order entry and are capable of doing so, NOE may be considered a beneficial alternative order entry method.

(J Med Internet Res 2010;12(1):e5) doi:10.2196/jmir.1284

KEYWORDS

Medical order entry systems; decision support systems, clinical; medication erors; Iran; infant, newborn; patient safety

Introduction

Medication errors can increase mortality and morbidity and add to healthcare costs [1]. Pediatric patients are at higher risk of medication errors because of weight-based dosing and difficulties in communicating with care providers [2]. Among all pediatric patients, neonates are the most vulnerable to medication errors because of their small body mass and extensive exposure to multiple medications in the neonatal ward or neonatal intensive care unit (NICU) [3]. Neonatal patients have special requirements, and during hospitalization, their weight and renal function may change frequently [4]. These changes demand frequent adjustment of prescription and administration dosages, which increases the risk of medication errors [5,6]. Dosing errors are the most prevalent type of errors in neonates, and most of these occur at the time of prescription [7]. Antibiotics are the most frequently prescribed type of drug involved in neonatal dosing errors [7,8]. Also reported have been severe adverse events due to miscalculated doses of anticonvulsants [9]. Therefore, strategies to prevent dosing errors of antibiotics and anticonvulsants in neonates should be prioritized.

In previous studies, computerized physician order entry (CPOE) with decision support functionalities has reduced dosing errors of antibiotics among inpatient neonates [10,11]. Despite promising results, only about 2% to 20% of the hospitals in high-income countries have successfully implemented CPOE [12]. Among several barriers to implementation, high implementation costs, physician resistance, and user frustration have been found to be the most important [13-15]. In many hospitals' order entry systems, nurses or other nonphysician health personnel enter medical orders into the computer [16]. Even in hospitals that have successfully implemented strictly physician order entry (POE), some orders are entered by the nurses [17]. Some investigations have shown that nurses often have more positive attitudes toward computerized systems than physicians [18]. Therefore, the involvement of nurses in the order entry process may increase the rate of success and reduce physicians' resistance [16,17]. In a number of recent studies, researchers have defined CPOE as computerized provider order entry that includes participation by credentialed nurses [19].

The successful implementation of POE becomes even more complicated in middle- and low-income countries with economic and human resource constraints [20]. One such country is the Islamic republic of Iran, a country in the Middle East with a population of 70 million as of 2006 [20,21]. Iran is cooperating with the World Health Organization to extend the use of information technology and evidence-based decision making in the health sector [22].

Studies performed in Iran demonstrate that medication dosing errors and adverse drug events (ADE) are significant problems for the Iranian healthcare system [23,24]. In almost all Iranian hospitals that have implemented electronic medical record systems, nurses or professional operators enter medical information into the computer. Physicians do not interact with the system at all, or their interaction is limited [20].

In 2007, a POE system was implemented in the neonatal ward of an Iranian teaching hospital. The aim of this project was to investigate whether the implementation of the system reduced medication errors and to investigate transferability of the system to other wards of this hospital as well as to other teaching hospitals in Iran [20]. The introduction of the system was found to reduce medication errors of antibiotics and anticonvulsants [25]. However, the busy residents were reluctant to enter all prescribed orders into the computer. After several interview sessions with attending physicians, residents, and nurses, a new implementation model was introduced to address this challenge.

In the new order entry model, nurses entered the orders into the computer, and the resident physicians verified the correctness of the orders and countersigned them electronically. Despite the successful implementation of this method, its effectiveness in reducing medication errors still needed to be examined.

The aim of this study was thus to determine whether the new collaborative order entry method was as effective as the strictly physician order entry method in reducing nonintercepted dose and frequency medication errors of antibiotics and anticonvulsants.

Methods

Setting

The study was conducted in the neonatal ward of a 400-bed tertiary care referral teaching hospital (Besat) in the capital city of Hamadan that provides a variety of clinical services. Hamadan is a province in the northwest of Iran with almost 1,700,000 inhabitants. Besat's neonatal ward is a 17-bed clinical ward that includes two NICU beds.

System description

Hospital Information System (HIS)

Sayan-HIS (Sayan Rayan Co Ltd, Hamadan, Iran) is a commercial patient-centered hospital information system (HIS) that is used in all fifteen university-affiliated hospitals in Hamadan. It is a client-server application that uses MS-SQL server 2003 as its database. Users interact with the system in a local area network and through desktop computers installed at workstations. The system includes an administrative as well as a clinical information system. The administrative information system handles patient billing and the insurance company interface as well as providing various reports for the financial controllers and management.

Clinical Information System

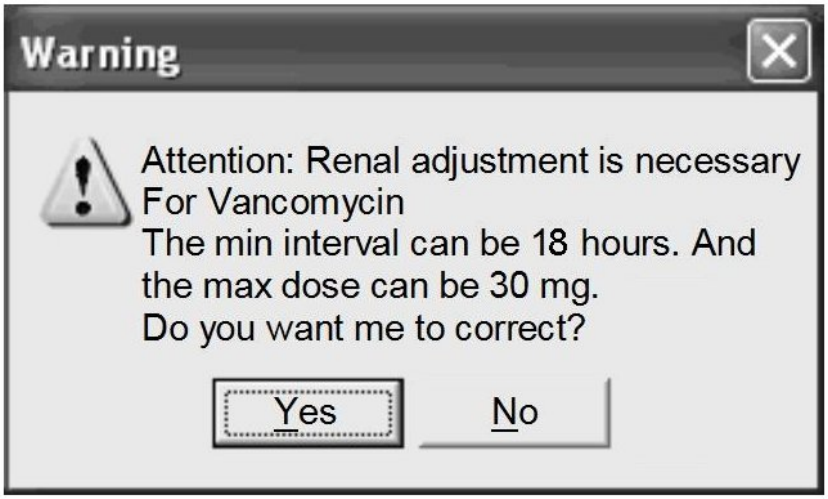
The clinical information system of Sayan-HIS includes an order-entry based prescription system. When the physician's orders are entered into the computer, the prescription system delivers the requested orders for medications, lab tests, and imaging to the relevant hospital sections at the appropriate time. The system limits the selection of drugs and their pharmaceutical forms (vial, ampoule, tablet, etc) through drop-down lists and preconstructed orders. The system was functional and routinely used with all explained features in all wards of the Besat hospital at the time of this study. The system also includes a rule-based clinical decision support system that is capable of alerting and correcting an erroneously prescribed dose or frequency of an antibiotic or anticonvulsant for neonatal patients.

Clinical Decision Support System (CDSS)

The dose and frequency decision support system was developed in 2007. The knowledge base was completed for all routine antibiotics and anticonvulsants by using the local guidelines of best practice based on pediatric reference books approved by the National Board of Pediatrics in Iran [26-29]. Prescription decision criteria were based on each patient's clinical diagnosis, age, weight, gestational age, and estimated glomerular filtration rate (GFR). Three neonatal specialists and one pediatric nephrologist reviewed and approved the CDSS calculation methods.

The system displayed warning messages on the prescription page whenever it detected a dose or frequency medication error based on the previously mentioned criteria (Figure 1). The warning supplied the appropriate dose and/or frequency as well as an explanation as to why the warning had appeared. The prescriber was then allowed to comply with the warning's recommended dosage or to ignore it. The responses of the prescribers to the warnings were recorded by the system in an error registration table. A detailed description of the CDSS and its interactions with the prescription system was presented in a previous paper [25].

Figure 1. A warning message for dose and frequency errors that gives the reason for the warning (Note that the figure shows a translated mockup)



Inclusion criteria and study population

The study population consisted of neonatal patients who were prescribed antibiotics for infectious diseases or anticonvulsants for seizure and who received at least one dose of these drugs. All orders for antibiotics and anticonvulsants for these patients were included.

Definition of medication errors

Normal ranges of doses and frequencies of the selected medications were calculated based on the published references cited above [26-29]. Medication errors for the purpose of this study were defined as overdoses or underdoses or curtailed or prolonged intervals.

In this study, we focused on both prescription and transcription errors but not on administration errors. A prescription error was defined as an error that occurred during the prescription stage. The prescription stage included errors in orders written initially on paper or directly entered into the computer by providers, or when a dose should have been changed but the prescriber ignored the computer warning or neglected to correct the prescription. The latter case mostly occurred when a dose decision criteria (age group, weight, GFR, etc) was changed during the hospitalization period.

A transcription error was defined as an error that occurred after the prescription stage. This type of error could have happened when information was transferred from handwritten orders to the computer or vice versa. This type of error could also have happened during registration of the doses in the paper-based medication administration chart or in the paper-based Kardex kept at the nurses' station that contains each patient's scheduled medications. A medication order with errors in both the prescription and the transcription stage was considered a prescription error.

Medication errors that did not reach patients were categorized as intercepted errors, and medication errors that reached patients were categorized as nonintercepted errors. Interception of medication errors could have occurred during two different phases in the prescription process. The first phase was when the physician prescribed an erroneous dosage but the CDSS corrected the error following a warning with which the physician complied. The second phase was when an erroneous dosage was detected and corrected by the nurse or physician before the medication was administered to the patient.

Nonintercepted medication errors could have occurred during the prescription phase or during the transcription phase. Such errors could have happened if the prescriber registered medications erroneously into the paper-based order or directly into the computer or if the paper-based order was erroneously transcribed into the system or Kardex by the nurse.

Data collection and review process

This study was a prospective study. Physicians' responses to the warnings were stored in a table together with both the erroneous and the corrected doses and frequencies. Therefore, it was possible to detect prescriptions that were initially incorrect but were intercepted by the warnings. In addition, one of the authors (AK) reviewed all relevant paper-based medical documents and electronic patient records. This included handwritten orders during the NOE period, electronic orders of both periods, and paper-based and electronic medication administration charts of both periods (Table 1).

Electronic orders and the electronic medication administration chart were tabulated automatically in the system's database during each period. However, data collection of the paper-based medication administration chart during both periods and the paper-based orders during the NOE period was performed after the NOE period (Figure 2). Analyses were performed after the completion of each period.

By triangulating different sources of data we could detect those medications that were prescribed erroneously and were not intercepted by the warnings, but were intercepted by physicians or nurses before the medications were administered to the neonates. In these cases, the electronic orders were registered with erroneous doses, but the paper-based medication administration charts were registered with the correct doses.

As well, medications that were prescribed erroneously and also registered in paper-based medication administration charts with erroneous doses were considered nonintercepted prescription errors. Medications that were prescribed with correct dosages but were registered in the paper-based medication administration chart with erroneous doses or frequencies were considered nonintercepted transcription errors. This last type of error could have occurred because of frequent transcriptions between paper-based and electronic orders, orders and the nursing Kardex, or the Kardex and paper-based medication administration charts.

Measuring medication errors

The rate of nonintercepted medication errors was calculated using the following four measures.

Patient-day was defined as one day of hospitalization for a patient who received medication therapy during the day. If all medications in all prescribed orders on the same day were correct, that day was as one correct patient-day, otherwise it was counted as an erroneous patient-day.

Medication-day was defined as a medication that was prescribed and continued for a patient on the same day. If all prescribed orders of a medication on the same day were correct, it was counted as one correct medication-day, otherwise it was counted as an erroneous medication-day.

Order was defined as a collection of prescribed medications, lab tests, imaging, and so on written by a physician for a patient during or after a visit to the patient’s bedside. If all prescribed medications in the same order were correct, it was counted as one correct order, otherwise it was counted as an erroneous order.

Ordered medication was defined as a medication prescribed in an order. If the prescribed medication was correct, it was counted as one correct ordered medication, otherwise it was counted as an erroneous ordered medication.

Study periods and their characteristics

This study was designed to compare two medication order entry methods each of which was studied over a 2-month periods. During period 1, or POE, physician order entry was followed by nurse verification and countersignature; during period 2, or NOE, nurse order entry was followed by physician verification and countersignature. The study was conducted between December 2007 and September 2008 (Table 1).

Table 1. Computerized order entry periods at the neonatal ward of the Besat hospital

	Period 1: Dec 2007 - Feb 2008	Period 2: Jul - Sep 2008
Intervention	POE[a]	NOE[b]
Order entry	Resident physicians	Nurses
Verification and countersignature	Nurses	Resident physicians
CDSS[c] functionality	Warnings	Warnings
When warnings displayed	Order entry	Countersignature
Documentation	E-Prints[d]	HWO[e]+E-Prints[d]
Review process	EO[f]+PBMAC[g]+EMAC[h]+ERT[i]	HWO[e]+EO[f]+PBMAC[g]+EMAC[h]+ERT[i]

[a] Physician order entry

[b] Nurse order entry

[c] Clinical decision support system

[d] Electronic prints of prescriptions

[e] Handwritten orders

[f] Electronic orders

[g] Paper-based medication administration chart

[h] Electronic medication administration chart

[i] Error registration table

Period 1: Physician Order Entry Followed by Nurse Verification and Countersignature (POE)

During period 1, resident physicians entered all prescription orders directly into the computer and paper-based orders were eliminated (Figure 3). To reduce possible data entry errors, a nurse verified and countersigned each electronic order that physicians had entered into the computer. This verification was designed to reduce the likelihood of making typographical errors or of selecting incorrect drugs from the drop-down menus. A further design consideration was to remind physicians about obvious dosing errors of those medications that were not included in the knowledge base (ie, drug groups other than antibiotics and anticonvulsants) and consequently warnings could not help to prevent them. Because Iranian law does not permit electronic signatures, each electronic order was printed and saved in the patient's medical file after it was countersigned [30] (Table 1).

Also in this period, each prescription line was assessed by the decision support system as it was prescribed by the resident physician. When a resident had ignored a warning, the ignored warning appeared each time the resident renewed the order with the same erroneous dose and frequency, or when the resident prescribed a new dosage that was also erroneous (Figure 2).

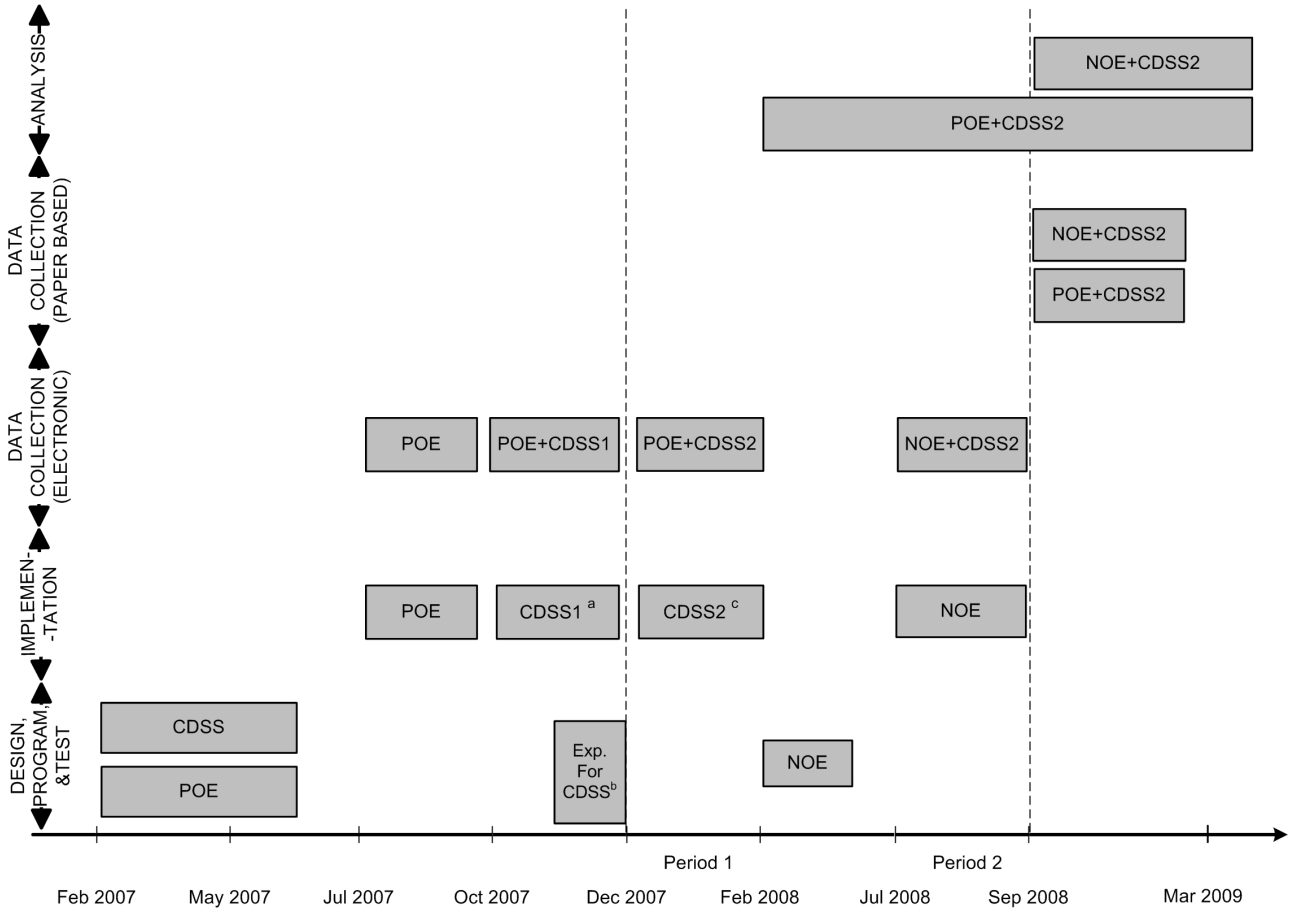
The design, programming, and testing of the decision support system for the POE method started in February 2007 (Figure 2). During this period, the functionality of the CDSS was gradually developed [25]. During period 1 of the current study, the frequency and format of the displayed warnings had been optimized, and these remained unchanged in period 2. Therefore, this period of POE was selected to be compared with NOE.

Period 2: Nurse Order Entry Followed by Physician Verification and Countersignature (NOE)

During period 2, the care providers of period 1 switched their roles in order entry and countersignature, vis-à-vis (Table 1 and Figure 3). Resident physicians wrote the initial orders on the prescription papers and delivered them to the nurses who subsequently entered them into the computer. The residents then verified and countersigned the orders electronically. Warnings appeared only at the time of physicians' countersignatures. Therefore, in this new model, warnings appeared to the residents but not to the nurses (Figure 3). This strategy was adopted because in a previous study of CPOE in Iran, physicians were reluctant to let their errors be disclosed to nurses and wished to receive the warnings themselves [20]. However, after the implementation of POE, the residents started to resist performing the order entry because they perceived it to be very time consuming. However, they still wanted to receive the warnings themselves without allowing the nurses to see them. The new model was designed in close collaboration with the involved physicians and nurses to address this issue.

After the physician's verification and countersignature, the electronic prescription was printed, and if a warning had been complied with that led to a change of dose or frequency, both the nursing Kardex and the patient file were updated (Figure 3). In period 2, both electronic prints and handwritten prescription papers were saved in the patient's file.

Figure 2. Development, implementation, and evaluation of clinical decision support system (CDSS), physician order entry (POE), and nurse order entry followed by physician confirmation (NOE) in the neonatal ward of the Besat hospital



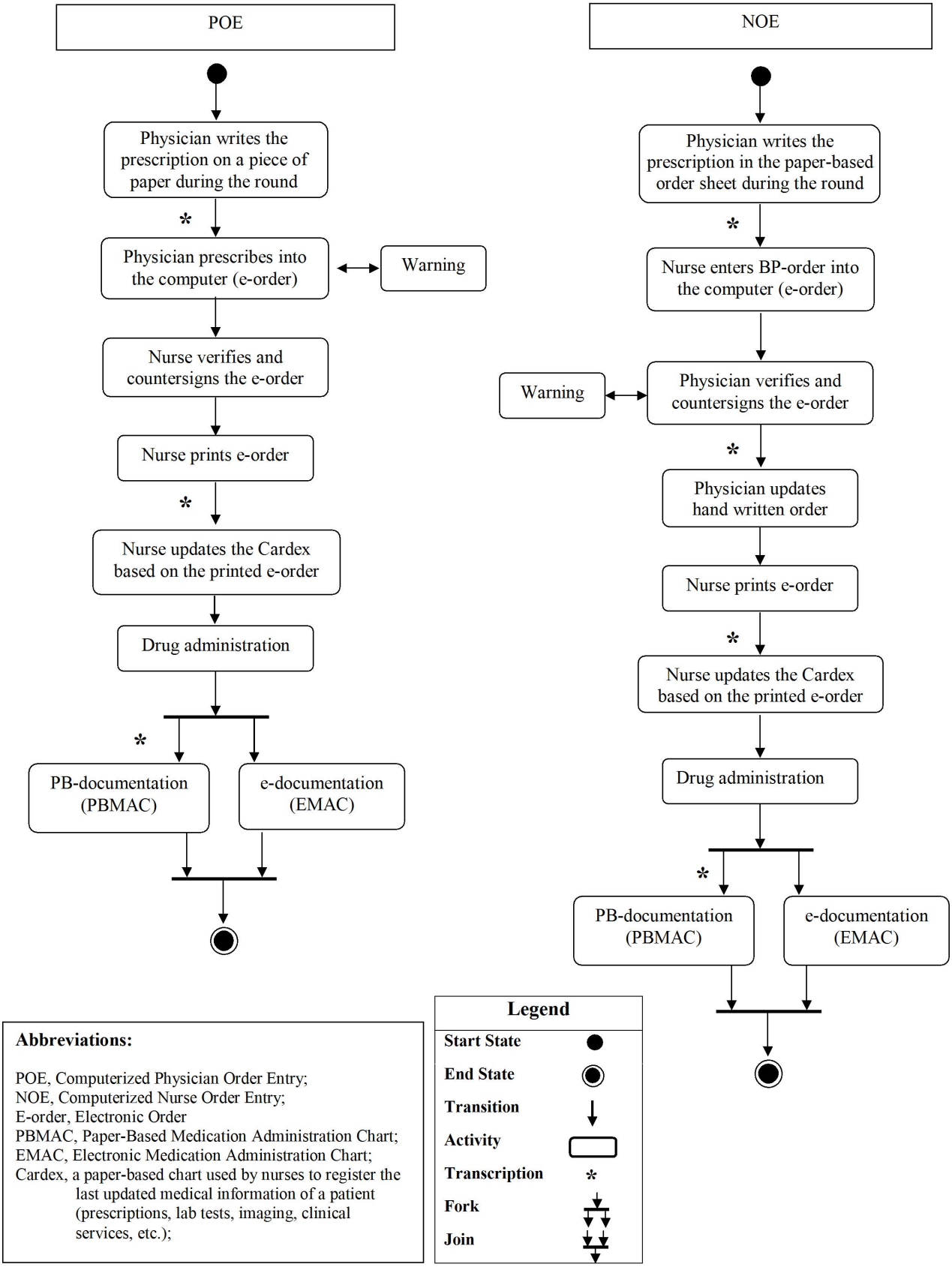
CDSS, clinical decision support system; POE, physician order entry method; NOE, nurse order entry followed by physician confirmation

[a] The warnings appeared at the time of an erroneous new prescription or change of the dosing criteria and provided no explanation

[b] An explanation was developed and added to the warning which could explain the reason that the alert appeared for

[c] The warnings appeared at each erroneous order and provided an explanation

Figure 3. Medication prescription and administration workflows during the POE and NOE periods in the neonatal ward of the Besat hospital



Statistics

Descriptive statistical analysis was used to determine medians for continuous variables and percentages for categorical variables. The median and the 25th and 75th percentiles of medication errors and the medians of age at admission, gestational age, and length of hospital stay were computed. Chi-square tests were performed for nonordinal categorical

variables. For continuous variables, the nonparametric Mann-Whitney tests were employed to determine differences in the median error rates between the POE and NOE periods when there was remarkable deviation from normality. Rates of errors were reported pertaining to orders, ordered medications, medication days, and patient days. Error rate differences between the POE and NOE periods were calculated as d = absolute value of the error rate during the POE period minus the error rate during the NOE period. Rate ratio (RR) was defined as the rate of errors during the NOE period divided by the rate of errors during the POE period. RR < 1 indicates that NOE has a "protective effect," and RR > 1 demonstrates that NOE has an "incremental effect" for medication errors. Confidence intervals for the ratios were determined under the assumption that the number of events per 100 patient-days followed a Poisson distribution. Miettinen's test-based approximation was used to calculate the confidence interval for the rate ratios. The level of statistical significance was specified at 0.05. Statistical analyses were performed using SPSS version 17 (SPSS Inc, Chicago, IL, USA). EPI Info version 6.0 (Centers for Disease Control and Prevention, Atlanta, GA, USA) was used to calculate chi-square for trend Mantel extension test [31] to examine an increasing or decreasing linear trend in the severity of overdose errors during the NOE period compared with the POE period.

Ethical considerations

The National Ethical Committee of the Ministry of Health and Medical Education in Iran granted permission for this study in 2005. All physicians and nurses who participated volunteered to take part in the study, and a verbal informed consent was obtained and tape recorded.

Results

A total of 158 neonates were included in this study (Table 2). No significant differences were observed in the distribution of sex, age at admission, or gestational age of these neonates between the POE and NOE periods.

Table 2. Distribution of the characteristics of patients included in the study, numbers of orders, and numbers of medications in the two study periods, POE and NOE

	POE	NOE
Patients	69	89
Male/female	35/34	41/48
Median age at admission (days)	7	5
Median gestational age (weeks)	38	38
Orders	972	978
Ordered medications[a]	2357	2297
Patient days[b]	601	648
Medication days [c]	1466	1492
Median length of hospital stay (days)	9.1	6.7

[a] A prescribed medication in an order is one ordered medication

[b]The number of days that patients received antibiotics or anticonvulsants

[c]The number of days that included medications were continued for patients

Medication errors were reduced to an equal extent during both the POE and NOE periods (Table 3). However, as the rate of errors that were intercepted by the warnings increased from 4.5% in the POE period to 8.1% in the NOE period (rate ratio 1.80, 95% confidence interval [CI] 1.43, 2.27; P < .001), the rate of nonintercepted errors dropped from 12.8% to 7.6% respectively (rate ratio 0.60, 95% CI 0.50, 0.71; P < .001). Most of the intercepted errors were caught by the warnings at the prescription stage; only a few errors were subsequently detected and intercepted by nurses or physicians before they were administered to the patients. The number of errors that were intercepted by the care providers was not significantly different between the two periods.

Table 3. Intercepted and nonintercepted medication errors and their rate ratio in the POE and NOE periods

Type of medication error	POE (n[a]=2357)	NOE (n[a]=2297)	RR[b] (95% CI[c])
Intercepted by the warnings	106 (4.5)[d]	186 (8.1)	1.80 (1.43, 2.27)[e]
Intercepted by care providers[f]	12 (0.5)	11 (0.5)	0.94 (0.42, 2.13)
Nonintercepted	301 (12.8)	175 (7.6)	0.60 (0.50, 0.71)[e]
Total	419 (17.8)	372 (16.2)	0.91 (0.8, 1.03)

[a] n = number of ordered medications

[b] Rate ratio

[c] Confidence interval

[d] Numbers in parentheses are percentages of errors calculated as [(number of errors)/ n] * 100

[e] $P < .001$

[f] Includes errors intercepted by nurses or physicians after the prescription stage and before the administration

Table 4 depicts different measurement units employed to calculate the rate and rate ratios of nonintercepted medication errors following the implementation of NOE in contrast to POE period. All measurements showed a highly significant reduction of medication errors from the POE period to the NOE period. However, the highest rate difference (9.5%) was seen when calculated according to patient days (rate ratio 0.61; 95% CI 0.49, 0.77; $P < .001$), and the lowest (5.2%) when using the ordered medications method (rate ratio 0.60; 95% CI 0.50, 0.71; $P < .001$). NOE showed a greater reduction effect on medication errors in all four calculation methods.

The median nonintercepted error per patient decreased from 2 (25th percentile = 0 and 75th percentile = 5) in the POE period to 0 (25th percentile = 0 and 75th percentile = 2) in the NOE period ($P = .005$). In the POE period, about 38% (26/69) of the patients did not experience any dosing errors, while in the NOE period, about 53% (47/89) of them were error-free (the rate difference was 15%).

Table 4. Rates and rate ratios of nonintercepted medication errors in POE and NOE using different measurements

Measurement unit	POE Errors/n (%)[a]	NOE Errors/n (%)[a]	RR[b] (95% CI[c])
Orders	221/972 (22.7)	142/978 (14.5)	0.64 (0.53, 0.77)[d]
Ordered medications	301/2357 (12.8)	175/2297 (7.6)	0.60 (0.50, 0.71)[d]
Medication-days	211/1466 (14.4)	129/1492 (8.6)	0.60 (0.49, 0.74)[d]
Patient-days	147/601 (24.5)	97/648 (15.0)	0.61 (0.49, 0.77)[d]

[a] errors is the number of errors per measurement unit; n is the number of measurement units (see Table 2); the number in parentheses is percentage of errors calculated as [(number of errors)/ n] * 100

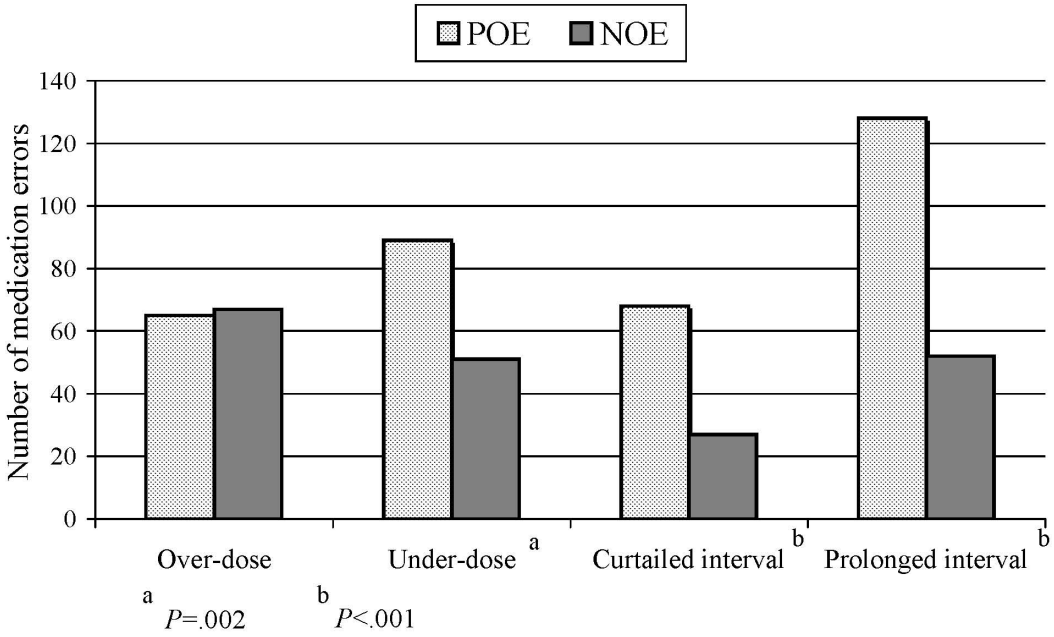
[b] Rate ratio

[c] Confidence interval

[d] $P < .001$

We divided nonintercepted medication errors into overdose, underdose, curtailed interval, and prolonged interval. We found that all subtypes of errors except overdose errors decreased significantly from the POE to the NOE period ($P = .002$ for underdose, and $P < .001$ for curtailed and prolonged interval errors) (Figure 4). The rate of overdose errors remained unchanged. However, there was a linear decreasing trend in severity of the overdose errors in the NOE period compared with the POE period (chi square for trend = 5.2; $P = .02$). The maximum registered overdose was less than 250% of the normal dose in the NOE period and less than 300% in the POE period. Two-fold or greater dosing errors occurred in about 25% (16/65) of overdosed medications in the POE period, while this occurred only in about 7% (5/67) of overdosed medications in the NOE period ($P = .007$).

Figure 4. Subtypes of nonintercepted dose and frequency medication errors



The rate of prescription errors decreased significantly from 10.3% in the POE period to 4.6% in the NOE period (rate ratio 0.45; 95% CI 0.36, 0.56; $P < .001$) (Table 5). Meanwhile, transcription errors showed a slight increase from 2.5% in the POE period to 3% in the NOE period. However, in both the POE and NOE periods, the majority of nonintercepted errors occurred in the prescription phase (80% in the POE and 60% in the NOE period). Therefore, the overall rate of errors also decreased from the POE period to the NOE period (rate ratio 0.60; 95% CI 0.50, 0.71; $P < .001$).

Table 5. Nonintercepted prescription and transcription errors in the ordered medications of POE and NOE and their rate ratio

Error type	POE (n[a] = 2357)	NOE (n[a] = 2297)	RR[b] (95% CI[c])
Prescription errors	242 (10.3)[d]	106 (4.6)	0.45 (0.36, 0.56)[e]
Transcription errors	59 (2.5)	69 (3.0)	1.20 (0.85, 1.69)
Total	301 (12.8)	175 (7.6)	0.60 (0.50, 0.71)[e]

[a] n = number of ordered medications

[b] Rate ratio

[c] Confidence interval

[d] Numbers in parentheses are percentages of errors calculated as [(number of errors)/ n] * 100

[e] $P < .001$

Many prescription errors occurred because the prescriber set an erroneous dose at the time of prescription. Other errors occurred when one or more of the dose decision criteria (age, weight, GFR, etc) had changed since the last prescription but the prescriber did not change the prescribed order and repeated the previously ordered dose and frequency (Table 6). In the NOE period, many transcription errors occurred when the electronic order was updated following a warning with which the prescriber had been complied, but the paper-based order was not updated or was updated with a different dose or frequency. This type of error did not happen in the POE period since the handwritten orders were eliminated in this period. The number of errors that occurred following incorrect registration of the paper-based medication administration chart although the electronic medication administration chart was correct, did not significantly differ between the POE and NOE periods. The rate of errors that occurred because the prescriber neglected to update the paper-based Kardex was approximately the same during the NOE and POE periods.

The total number of warnings was 312 in the POE period and 339 in the NOE period. The number of warnings with which the prescribers complied increased significantly from 44% (136/312) in the POE period to 68% (232/339) in the NOE period ($P < .001$).

Table 6. Distribution of nonintercepted medication errors at different registration steps of the POE and NOE periods

Reasons for dose and frequency errors	P/T[a]	POE (n[b]=2357)	NOE (n[b]=2297)	RR[c] (95% CI[d])
Ordered dosage was initially incorrect	P	163 (6.9)[e]	70 (3.0)	0.44 (0.34, 0.58)[f]
Order continued with the previous dose despite the change in dosing criteria	P	79 (3.4)	36 (1.5)	0.47 (0.32, 0.69)[f]
PB-order inconsistent with E-order	T	0 (0.0)	22 (1.0)	N/A[f]
PBMAC inconsistent with EMAC[g]	T	24 (1.0)	13 (0.6)	0.56 (0.28, 1.09)
Prescribed order changed but still the previous dose administered (Kardex was not updated)	T	35 (1.5)	34 (1.5)	1.00 (0.62, 1.59)
Total		301 (12.8)	175 (7.6)	0.60 (0.50, 0.71)[f]

[a] P/T: prescription or transcription error

[b] n = number of ordered medications

[c] Rate ratio

[d] Confidence interval

[e] Numbers in parentheses are percentages of errors calculated as [(number of errors)/ n] * 100

[f] $P < .001$

[g] PBMAC, paper-based medication administration chart; EMAC, electronic medication administration chart

Discussion

Previous studies have highlighted a low compliance with POE and a high resistance to acceptance of it among physicians, as well as the failure of POE systems in developed countries [12-14,32]. The initial intention of this study was to investigate whether NOE as an alternative order entry method was at least as effective as POE in reducing medication dosing errors. Surprisingly, we observed that the overall rate of nonintercepted dose and frequency medication errors was in fact lower under NOE than POE.

One reason for the lower error rate is that the prescribers complied with a higher rate of warnings in the NOE than in the POE period. The result was a significant reduction in the rate of nonintercepted prescription errors. Other studies have also reported that decision support systems can reduce prescription errors if prescribers comply with the system's recommendations [33,34]. Since in the POE period a majority of the nonintercepted errors occurred in the prescription stage, reduction of prescription errors resulted in an overall reduction of nonintercepted errors. Previous studies in the pediatrics and neonatal settings show that a majority of errors occur in the prescription stage [7].

In addition, most of the errors that were not intercepted by the warnings in the prescription stage reached the patients. Only a few of these errors were caught by the care providers. This reveals the importance of the dose decision support system and prescribers' compliance with the system's recommendations in this context. In developed countries, in addition to decision support systems, clinical pharmacists in many hospitals interact with care providers and supervise the preparation and administration of medications. In many cases, the pharmacy department is responsible for preparing ready-to-administer doses. The results of two studies, one in the United States and one in the United Kingdom, demonstrated a 66% to 80% reduction of medication errors following the active involvement of a senior clinical pharmacist in the clinical rounds [35,36]. However, in most hospitals in Iran, pharmacists and clinical pharmacologists do not participate in clinical rounds. The pharmacy does not prepare ready-to-administer doses; nurses in the wards are responsible for these. In Iranian hospitals, many responsibilities are left to the nurses. This is mostly because a very hierarchical system exists in these hospitals [20]. Hospital managers often assign to nurses, who are at the bottom of this hierarchy, tasks that physicians or pharmacists object to performing [20]. Medical data entry is one of these tasks. In Iranian hospitals, there are few legal or administrative incentives for physicians to enter medical data into electronic systems [20,37]. Therefore, strategies such as NOE, which require less physician time, may increase physicians' compliance and result in a more sustainable implementation of computerized provider order entry systems.

In addition, there are several other possible explanations for increased compliance in the NOE period. One explanation is that in the strictly physician order entry period, resident physicians were more likely to have focused on data entry than on the warnings. They may have ignored the warnings

unintentionally because of frustration and stress following a prolonged data entry session. A previous study showed that it is difficult to successfully implement systems that physicians consider to be time consuming [14]. The authors stated that prolonged data entry and user frustration were important causes of the failure of CPOE in their study [14]. However, in the nurse order entry method, physicians needed only to focus on prescription errors and warnings. This could have increased their attention to the displayed warnings and resulted in better compliance. It is also possible that the new collaborative environment in the NOE period created a better understanding of the advantages of the CDSS and resulted in better physician compliance with the system's recommendations. Today, more and more hospitals in western countries are attempting to redefine traditional borders between doctors and nurses by creating closer collaboration between them in all clinical activities [38,39]. In countries like Iran, where a hierarchical and physician-centered atmosphere exists in clinical settings [20], for CPOE systems to be successful, it is important that managers and policy makers create a collaborative and patient-centered climate.

Another possible explanation for higher compliance in the NOE period, is that NOE was designed in close collaboration with care providers and reflected their opinions. Therefore, care providers were more compliant with the new order entry method. As other studies have emphasized, care providers' acceptance and their collaboration in the development process are key factors in successful implementation of computerized order entry systems [17].

In addition, the reduction in medication errors can also be attributed to the fact that prescription orders may have been double-checked by the prescribing physicians in the NOE period. In the NOE model, prescribers had to check transcribed orders before signing them. This provided them with the possibility of double-checking what they had already prescribed before they received any warnings. This double-checking, independent of CDSS warnings, can also explain the observed reduction in prescription error rates in the NOE period.

In our study, the increase of transcription errors from the POE to the NOE period was small. Considering the workflow of the two order entry methods (Figure 3), NOE seems to be more complex than POE. Therefore, the rate of transcription errors should be higher in the NOE model than the POE model. However, since Iranian law prevents elimination of paper-based medical records [30], any medical order entry system, even POE, includes redundant recordings and documentation. Therefore, in such a context, POE has no apparent advantage over NOE in terms of transcription errors. In the United States and some European countries, where computerized order entry has reduced paperwork, POE has become a powerful tool to prevent transcription errors [33,40-42].

In our study, despite the nonsignificant difference in the overall rate of transcription errors between the POE and NOE periods, there are certain types of these errors that could be eliminated by POE. When a physician directly prescribes into the computer and prints the order, there can be no discrepancy between the electronic and paper-based order. In contrast, when using NOE, a physician must write a paper-based order and sign it for the nurse so that the nurse can enter the order into the computer. Since this paper-based order is a legal document, when a warning has been accepted, the resident must also update the paper-based order; negligence may result in nonintercepted transcription errors as in our study. Other types of transcription errors were not significantly different between the POE and NOE periods because after the prescription stage, the transcription and administration flows are the same in both systems.

In order to reduce transcription errors in Iran, prescription workflow should be simplified, and paper work should be limited. These strategies can save time, reduce costs, and may directly affect care providers' satisfaction resulting in higher acceptance. In Iran, many care providers complain that paperwork has dominated clinical care and that computerized systems have created many redundant registrations and documentation [20]. However, adapting Iranian law to demands from the digitized world is a challenge.

Methodological considerations

In this study, we calculated the number of nonintercepted medication errors using four measurement methods. Previous studies have used one or more of these measurement methods to report medication errors [43,44]. However, the calculation method affects the error rate considerably. For example, in our study, medication error rates reported per patient-day were twice the error rates reported per ordered medication. The reason is that in the patient-day method, several medications in several orders on the same day were counted as one unit. Therefore, if even one of these medications was erroneous, that patient-day was counted as erroneous. In the medication-day method, medications were analyzed separately, but a medication that was repeated in several orders on the same day was counted as one medication-day. If a medication was erroneous in one of these orders, then that medication-day was counted as erroneous. Reporting errors per order solved the problem of putting several orders in the same package; however, simultaneously prescribed medications in an order became one unit of analysis. Therefore, if one of these medications was erroneous, then that order was considered erroneous. In the ordered-medication method, each medication in each order was the unit of analysis. Therefore, an erroneous medication in one order did not adversely affect the other ordered medications.

In addition to the error calculation method, the data collection method and review process can affect the error rate [43]. Studies such as Simpson et al [36] that are based on critical or spontaneous reports can detect only a fraction of medication errors [44]. The reason is that these methods are heavily dependent on the individuals and their willingness to share their errors. In hospitals where staff members are afraid of punishment, there will be a lower tendency to openly report critical errors. In the study by Simpson et al [36], the error rate before the educational intervention by a pharmacist was 24.1 per 1000 neonatal days and after the intervention was 5.1 per 1000 neonatal days.

Chart reviews, especially when they are coupled with voluntary reports as in the study conducted by Kaushal et al [7], can detect

a higher proportion of prescription errors; the error rate in this study was 5.5 per 100 orders. Direct observation is appropriate for detecting administration errors [44], although it is prone to biases such as the Hawthorne effect [45]. Furthermore, studies like Cordero et al [10] that have reviewed handwritten and electronic medical records have detected a higher rate of medication errors. Such studies have reported the error rate to be as high as 13 per 100 orders.

In our investigation, we reviewed both the handwritten and electronic medical records of orders and nursing charts in both periods. We found the rate of nonintercepted errors to have been 22.7 and 14.5 per 100 orders in the POE and NOE periods respectively. The error rate was 245 per 1000 patient-days in the POE period and 150 per 1000 patient-days in the NOE period.

In summary, methods for calculating and reporting medication errors in neonatal settings are diverse and the results difficult to compare. However, based on all four calculation methods, NOE resulted in a lower rate of nonintercepted dose and frequency medication errors than POE.

Limitations

This study has a number of limitations. The study was performed in a neonatal setting; therefore, the results may not be generalizable to adults. We selected to study prescribing for the patient group at two points in time because we could not divide patients into two groups—a study group and a control group—in the neonatal ward. Implementation of different medical order entry systems poses a systemic change of the prescription flow in the ward. Moreover, we could not form a control group from another ward of the hospital since the guidelines and dose calculation criteria were very different between the neonatal and other wards.

Since the residents were still in training, their knowledge would be expected to increase over time. This can be a competing explanation of the findings, though previous studies have reported that dose calculation skill among pediatric residents is not related to their experience, grade, level of training, or commitment to recheck their calculated doses [46,47].

Additionally, the care providers knew that they were being studied. Therefore, they might have improved their performance during the study period, which could have led to the Hawthorne effect [45]. This could have affected the results in several ways. Residents knew that one of the purposes of the project was to find the appropriate medical order entry method and to extend this to the other wards of the hospital. It is possible that residents performed better in the NOE period to convince the hospital and university authorities to continue this method and not to return to POE. An opposite attempt by the nurses could also explain the high rate of transcription errors in the NOE period. However, the researchers could not find any evidence of such attempts.

Although the functionality of the decision support system was the same in the two periods, the residents' trust in the decision support system's functionality might also have increased over time. This could have led to a higher compliance among the prescribers in the NOE period and could have resulted in better prevention of medication errors in this period. An increase in trust could have happened because of positive experiences of prescribers or other care providers with the system over time and the sharing of those experiences with the others. It could also have happened because of a gradual increase in attention to patient safety among caregivers. However, the influence of these factors would be expected following positive experiences with decision support systems.

Conclusions

Since physicians' interaction with a dose decision support system is crucial in reducing medication errors, when physicians are resistant to entering orders into the computer and nurses are cooperative and capable of doing so, nurses can perform the order entry with physicians addressing the warnings. This strategy may significantly reduce physicians' resistance and increase their compliance with the system's recommendations. The new order entry method (NOE) can reduce nonintercepted medication dosing errors and increase safety among neonates as effectively as, or even better than, a strictly physician order entry method. However, NOE can increase transcription activities and paper work and add to the complexity of the prescription workflow. An ideal model should reduce physician's data entry workload without increasing the complexity. However, in countries like Iran, where elimination of paper-based medical documentation is not legally possible, POE has no significant advantage over NOE in reducing transcription errors. Therefore, in such settings, NOE could be considered to be a beneficial alternative order entry method.

Acknowledgments

The authors appreciate the executive and scientific support from the chair of pediatrics, Dr. Nasrollah Pezeshki; the head of the neonatal ward of the Besat hospital, Dr Fatemeh Eghbalian; and the scientific support of Dr. Behnaz Basiri, a faculty member of the Pediatrics Department of Hamadan University of Medical Sciences (HUMS). We are very thankful for the scientific contribution, constructive feedback, and the thorough review given by Professor Andrew Gaffney from the department of Cardiology of Vanderbilt University.

This study was supported by grants from the World Bank; the National Public Health Management Centre (NPMC), Tabriz, Iran; and the Hamadan University of Medical Sciences and Health Services, Hamadan, Iran. The sponsors had no involvement in the design, data collection, analysis, interpretation of data, writing of the manuscript, or the decision to submit the manuscript for publication.

Conflicts of Interest

None of the authors have any financial relationship with the local company who developed the system. The authors do not see any conflicts of interests in this study.

References

1. Classen DC, Pestotnik SL, Evans RS, Lloyd JF, Burke JP. Adverse drug events in hospitalized patients. Excess length of stay, extra costs, and attributable mortality. JAMA 1997;277(4):301-306. [Medline: 9002492] [doi: 10.1001/jama.277.4.301]
2. Levine SR, Cohen MR, Blanchard NR, Frederico F, Magelli M, Lomax C, et al. Guidelines for preventing medication errors in pediatrics. J Pediatr Pharmacol Ther 2001;6:426-442.
3. Kunac DL, Reith DM. Identification of priorities for medication safety in neonatal intensive care. Drug Saf 2005;28(3):251-261. [Medline: 15733029] [doi: 10.2165/00002018-200528030-00006]
4. Gray JE, Goldmann DA. Medication errors in the neonatal intensive care unit: special patients, unique issues. Arch Dis Child Fetal Neonatal Ed 2004 Nov;89(6):F472-F473 [FREE Full text] [Medline: 15499131] [doi: 10.1136/adc.2003.046060]
5. Daschner M. Drug dosage in children with reduced renal function. Pediatr Nephrol 2005 Dec;20(12):1675-1686. [Medline: 16133064] [doi: 10.1007/s00467-005-1922-9]
6. Patrick CH. Therapeutic drug monitoring in neonates. Neonatal Netw 1995 Mar;14(2):21-26. [Medline: 7746237]
7. Kaushal R, Bates DW, Landrigan C, McKenna KJ, Clapp MD, Federico F, et al. Medication errors and adverse drug events in pediatric inpatients. JAMA 2001 Apr 25;285(16):2114-2120 [FREE Full text] [Medline: 11311101] [doi: 10.1001/jama.285.16.2114]
8. Ross C, Banchy P. The key to CPOE. Health Manag Technol 2007 Nov;28(11):22, 24. [Medline: 18019863]
9. Lowry JA, Vandover JC, DeGreeff J, Scalzo AJ. Unusual presentation of iatrogenic phenytoin toxicity in a newborn. J Med Toxicol 2005 Dec;1(1):26-29. [Medline: 18072100] [doi: 10.1007/BF03160902]
10. Cordero L, Kuehn L, Kumar RR, Mekhjian HS. Impact of computerized physician order entry on clinical practice in a newborn intensive care unit. J Perinatol 2004 Feb;24(2):88-93. [Medline: 14872207] [doi: 10.1038/sj.jp.7211000]
11. Gard JW, Starnes HM, Morrow EL, Sanchez PJ, Perlman JM. Reducing antimicrobial dosing errors in a neonatal intensive care unit. Am J Health Syst Pharm 1995 Jul 15;52(14):1512-1513. [Medline: 7552892]
12. Aarts J, Koppel R. Implementation of computerized physician order entry in seven countries. Health Aff (Millwood) 2009;28(2):404-414. [Medline: 19275996] [doi: 10.1377/hlthaff.28.2.404]
13. Kaushal R, Jha AK, Franz C, Glaser J, Shetty KD, Jaggi T, et al; Brigham and Women's Hospital CPOE Working Group. Return on investment for a computerized physician order entry system. J Am Med Inform Assoc 2006;13(3):261-266 [FREE Full text] [Medline: 16501178] [doi: 10.1197/jamia.M1984]
14. Massaro TA. Introducing physician order entry at a major academic medical center: II. Impact on medical education. Acad Med 1993 Jan;68(1):25-30. [Medline: 8447888]
15. Poon EG, Blumenthal D, Jaggi T, Honour MM, Bates DW, Kaushal R. Overcoming barriers to adopting and implementing computerized physician order entry systems in U.S. hospitals. Health Aff (Millwood) 2004;23(4):184-190. [Medline: 15318579] [doi: 10.1377/hlthaff.23.4.184]
16. Roggow DJ, Solie CJ, Tracy MF, Gjere N. Clinical nurse specialist leadership in computerized provider order entry design. Clin Nurse Spec 2005;19(4):209-214. [Medline: 16027551] [doi: 10.1097/00002800-200507000-00013]
17. Ahmad A, Teater P, Bentley TD, Kuehn L, Kumar RR, Thomas A, et al. Key attributes of a successful physician order entry system implementation in a multi-hospital environment. J Am Med Inform Assoc 2002;9(1):16-24 [FREE Full text] [Medline: 11751800]
18. Callen J, Braithwaite J, Westbrook J. Differences in doctors' and nurses' assessments of hospital culture and their views about computerised order entry systems. Stud Health Technol Inform 2008;136:15-20. [Medline: 18487701]
19. Gurwitz JH, Field TS, Rochon P, Judge J, Harrold LR, Bell CM, et al. Effect of computerized provider order entry with clinical decision support on adverse drug events in the long-term care setting. J Am Geriatr Soc 2008 Dec;56(12):2225-2233. [Medline: 19093922] [doi: 10.1111/j.1532-5415.2008.02004.x]
20. Kazemi A, Ellenius J, Tofighi S, Salehi A, Eghbalian F, Fors UG. CPOE in Iran--a viable prospect? Physicians' opinions on using CPOE in an Iranian teaching hospital. Int J Med Inform 2009 Mar;78(3):199-207. [Medline: 18760960] [doi: 10.1016/j.ijmedinf.2008.07.004]
21. ; World Health Organization. World Health Statistics 2008. Geneva, Switzerland: WHO Press; 2008.
22. ; World Health Organization. Country Cooperation Strategy for WHO and the Islamic Republic of Iran 2005-2009. Geneva, Switzerland: World Health Organization; 2006.
23. Gholami K, Shalviri G. Factors associated with preventability, predictability, and severity of adverse drug reactions. Ann Pharmacother 1999 Feb;33(2):236-240. [Medline: 10084421] [doi: 10.1345/aph.17440]
24. Zargarzadeh AH, Emami MH, Hosseini F. Drug-related hospital admissions in a generic pharmaceutical system. Clin Exp Pharmacol Physiol 2007;34(5-6):494-498. [Medline: 17439421] [doi: 10.1111/j.1440-1681.2007.04600.x]

25. Kazemi A, Ellenius J, Pourasghar F, Tofighi S, Salehi A, Amanati A, et al. The effect of computerized physician order entry and decision support system on medication errors in the neonatal ward: Experiences from an Iranian teaching hospital. J Med Syst 2009 Jul. [doi: 10.1007/s10916-009-9338-x.]
26. Davis ID, Avner ED. Glomerular Disease. In: Behrman RE, Kleigman RM, Jenson HB, editors. Nelson Textbook of Pediatrics. 17th edition. Philadelphia, PA: WB Saunders; 2004:1733.
27. Gal P, Reed MD. Laboratory medicine, drug therapy, and reference tables- Medications. In: Behrman RE, Kleigman RM, Jenson HB, editors. Nelson Textbook of Pediatrics. 17th edition. Philadelphia, PA: WB Saunders; 2004:2473-2474.
28. Omokaro SO. Nephrology- normal values of GFR. In: Robertson J, Shilkofski N, editors. The Harriet Lane Handbook: A Manual for Pediatric House Officers. 17th edition. Philadelphia, PA: Elsevier Mosby; 2005:495.
29. Stoll BJ. Infections of the neonatal infant- Suggested dosage schedules for antibiotics used in newborns. In: Kleigman RM, Behrman RE, Jenson HB, Stanton BF, editors. Nelson Textbook of Pediatrics. 18th edition. Philadelphia, PA: Saunders; 2007:808.
30. Pourasghar F, Malekafzali H, Kazemi A, Ellenius J, Fors U. What they fill in today, may not be useful tomorrow: lessons learned from studying Medical Records at the Women hospital in Tabriz, Iran. BMC Public Health 2008;8(1):139 [FREE Full text] [Medline: 18439311] [doi: 10.1186/1471-2458-8-139]
31. Dean AG, Dean JA, Coulombier D, Brendel KA, Smith DC, Burton AH, et al. Epi InfoVersion 6.04a, a word-processing, database, and statistics program for public health on IBM-compatible microcomputers. Atlanta, GA: Centers for Disease Control and Prevention; 1996.
32. Poon EG, Blumenthal D, Jaggi T, Honour MM, Bates DW, Kaushal R. Overcoming barriers to adopting and implementing computerized physician order entry systems in U.S. hospitals. Health Aff (Millwood) 2004;23(4):184-190. [Medline: 15318579] [doi: 10.1377/hlthaff.23.4.184]
33. Mekhjian HS, Kumar RR, Kuehn L, Bentley TD, Teater P, Thomas A, et al. Immediate benefits realized following implementation of physician order entry at an academic medical center. J Am Med Inform Assoc 2002;9(5):529-539 [FREE Full text] [Medline: 12223505] [doi: 10.1197/jamia.M1038]
34. Eslami S, de Keizer NF, Abu-Hanna A. The impact of computerized physician medication order entry in hospitalized patients--a systematic review. Int J Med Inform 2008 Jun;77(6):365-376. [Medline: 18023611] [doi: 10.1016/j.ijmedinf.2007.10.001]
35. Leape LL, Cullen DJ, Clapp MD, Burdick E, Demonaco HJ, Erickson JI, et al. Pharmacist participation on physician rounds and adverse drug events in the intensive care unit. JAMA 1999 Jul 21;282(3):267-270 [FREE Full text] [Medline: 10422996] [doi: 10.1001/jama.282.3.267]
36. Simpson JH, Lynch R, Grant J, Alroomi L. Reducing medication errors in the neonatal intensive care unit. Arch Dis Child Fetal Neonatal Ed 2004 Nov;89(6):F480-F482 [FREE Full text] [Medline: 15499135] [doi: 10.1136/adc.2003.044438]
37. Pourasghar F, Malekafzali H, Koch S, Fors U. Factors influencing the quality of medical documentation when a paper-based medical records system is replaced with an electronic medical records system: an Iranian case study. Int J Technol Assess Health Care 2008;24(4):445-451. [Medline: 18828939] [doi: 10.1017/S0266462308080586]
38. Aarts J, Ash J, Berg M. Extending the understanding of computerized physician order entry: implications for professional collaboration, workflow and quality of care. Int J Med Inform 2007 Jun;76(Suppl 1):S4-13. [Medline: 16798068] [doi: 10.1016/j.ijmedinf.2006.05.009]
39. Dempsey C, Larson KW. Can't we all just get along? Know your role in fostering hospital-provider collaboration. Nurs Manage 2004 Nov;35(11):32-35. [Medline: 15543041]
40. Barron WM, Reed RL, Forsythe S, Hecht D, Glen J, Murphy B, et al. Implementing computerized provider order entry with an existing clinical information system. Jt Comm J Qual Patient Saf 2006 Sep;32(9):506-516. [Medline: 17987874]
41. Seeley CE, Nicewander D, Page R, Dysert PA. A baseline study of medication error rates at Baylor University Medical Center in preparation for implementation of a computerized physician order entry system. Proc (Bayl Univ Med Cent) 2004 Jul;17(3):357-61; discussion 361-2 [FREE Full text] [Medline: 16200120]
42. Shulman R, Singer M, Goldstone J, Bellingan G. Medication errors: a prospective cohort study of hand-written and computerised physician order entry in the intensive care unit. Crit Care 2005 Oct 5;9(5):R516-R521 [FREE Full text] [Medline: 16277713] [doi: 10.1186/cc3793]
43. Ghaleb MA, Barber N, Franklin BD, Yeung VWS, Khaki ZF, Wong ICK. Systematic review of medication errors in pediatric patients. Ann Pharmacother 2006 Oct;40(10):1766-1776. [Medline: 16985096] [doi: 10.1345/aph.1G717]
44. Flynn EA, Barker KN, Pepper GA, Bates DW, Mikeal RL. Comparison of methods for detecting medication errors in 36 hospitals and skilled-nursing facilities. Am J Health Syst Pharm 2002 Mar 1;59(5):436-446. [Medline: 11887410]
45. Hsueh Y. The Hawthorne experiments and the introduction of Jean Piaget in American industrial psychology, 1929-1932. Hist Psychol 2002 May;5(2):163-189. [Medline: 12096759] [doi: 10.1037/1093-4510.5.2.163]
46. Glover ML, Sussmane JB. Assessing pediatrics residents' mathematical skills for prescribing medication: a need for improved training. Acad Med 2002 Oct;77(10):1007-1010. [Medline: 12377676] [doi: 10.1097/00001888-200210000-00013]
47. Rowe C, Koren T, Koren G. Errors by paediatric residents in calculating drug doses. Arch Dis Child 1998 Jul;79(1):56-58 [FREE Full text] [Medline: 9771254] [doi: 10.1136/adc.79.1.56]

Abbreviations

ADE: adverse drug events
CDSS: clinical decision support system
CPOE: computerized physician order entry
EMAC: electronic medication administration chart
EO: electronic order
ERT: error registration table
GFR: glomerular filtration rate
HIS: hospital information system
HWO: handwritten order
NICU: neonatal intensive care unit
NOE: nurse order entry (followed by physician's verification and countersignature)
PBMAC: paper-based medication administration chart
POE: physician order entry
RR: rate ratio

Edited by G Eysenbach; submitted 13.06.09; peer-reviewed by H Pirnejad, N Creswick; comments to author 06.07.09; revised version received 09.12.09; accepted 09.12.09; published 26.02.10.

Please cite as:
Kazemi A, Fors UGH, Tofighi S, Tessma M, Ellenius J
Physician Order Entry Or Nurse Order Entry? Comparison of Two Implementation Strategies for a Computerized Order Entry System Aimed at Reducing Dosing Medication Errors
J Med Internet Res 2010;12(1):e5
URL: http://www.jmir.org/2010/1/e5/
doi:10.2196/jmir.1284
PMID:20185400

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org