Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 17 (2015), Issue 12 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Corrigenda and Addenda

Original Paper

An Online Health Prevention Intervention for Youth with Addicted or Mentally Ill Parents: Experiences and Perspectives of Participants and Providers from a Randomized Controlled Trial

Marla Woolderink[1,2], MPhil; Jill APM Bindels[1,2], PhD; Silvia MAA Evers[2,3,4], PhD, LLM; Aggie TG Paulus[2,3], PhD; Antoinette DI van Asselt[5,6], PhD; Onno CP van Schayck[2,7], PhD

[1]Clinical Epidemiology and Technology Assessment, Maastricht University Medical Center, Maastricht, Netherlands
[2]CAPHRI-School for Public Health and Primary Care, Maastricht, Netherlands
[3]Department of Health Services Research, Maastricht University, Maastricht, Netherlands
[4]Institute of Mental Health and Addiction, Trimbos Instituut, Utrecht, Netherlands
[5]Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, Netherlands
[6]Department of Pharmacy, University of Groningen, Groningen, Netherlands
[7]Department of General Practice, Maastricht University, Maastricht, Netherlands

Corresponding Author:
Marla Woolderink, MPhil
Clinical Epidemiology and Technology Assessment
Maastricht University Medical Center
Oxfordlaan 10
Maastricht, PO BOX 5800, 6202 AZ
Netherlands
Phone: 31 0031 043 3875
Fax: 31 0031 043 3874
Email: m.woolderink@maastrichtuniversity.nl

Abstract

Background: Mental illnesses affect many people around the world, either directly or indirectly. Families of persons suffering from mental illness or addiction suffer too, especially their children. In the Netherlands, 864,000 parents meet the diagnostic criteria for a mental illness or addiction. Evidence shows that offspring of mentally ill or addicted parents are at risk for developing mental disorders or illnesses themselves. The Kopstoring course is an online 8-week group course with supervision by 2 trained psychologists or social workers, aimed to prevent behavioral and psychological problems for children (aged 16 to 25 years) of parents with mental health problems or addictions. The course addresses themes such as roles in the family and mastery skills. An online randomized controlled trial (RCT) was conducted to assess the effectiveness of the Kopstoring course.

Objective: The aim was to gain knowledge about expectations, experiences, and perspectives of participants and providers of the online Kopstoring course.

Methods: A process evaluation was performed to evaluate the online delivery of Kopstoring and the experiences and perspectives of participants and providers of Kopstoring. Interviews were performed with members from both groups. Participants were drawn from a sample from the Kopstoring RCT.

Results: Thirteen participants and 4 providers were interviewed. Five main themes emerged from these interviews: background, the requirements for the intervention, experience with the intervention, technical aspects, and research aspects. Overall, participants and providers found the intervention to be valuable because it was online; therefore, protecting their anonymity was considered a key component. Most barriers existed in the technical sphere. Additional barriers existed with conducting the RCT, namely gathering informed consent and gathering parental consent in the case of minors.

Conclusions: This study provides valuable insight into participants' and providers' experiences and expectations with the online preventive intervention Kopstoring. It also sheds light on the process of the online provision of Kopstoring and the accompanying RCT. The findings of this study may partly explain dropout rates when delivering online interventions. The change in the (financial) structure of the youth mental health care system in the Netherlands has financial implications for the delivery of prevention programs for youth. Lastly, there are few RCTs that assess the effectiveness and cost-effectiveness of online prevention programs in the field of (youth) mental health care and not many process evaluations of these programs exist. This hampers a good comparison between online interventions and the expectations and experiences of the participants and providers.

Trial Registration: Nederlands Trial Register: NTR1982; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=1982 (Archived by WebCite® at http://www.webcitation.org/6d8xYDQbB)

(J Med Internet Res 2015;17(12):e274) doi:10.2196/jmir.4817

KEYWORDS

online-delivered course; process assessment; qualitative research; mental health; prevention; adolescents

Introduction

Mental illnesses affect many people around the world, either directly or through another person. The prevalence of mental illnesses, such as depression, is high [1] and many among those who suffer from mental illness or addiction are parents. The families of persons suffering from mental illness experience a degree of burden too [2,3]. Results of a Canadian survey show that 1 in every 10 children lives with a parent with a psychiatric disorder and 1 in every 6 children lives in a household with at least 1 person affected by a psychiatric disorder [4]. In these situations, mental illness is not only a problem for the patients, but also for their family and their children in particular. Various studies report that offspring of mentally ill/addicted parents are regarded at risk for developing mental illness (eg, depression and anxiety disorders) themselves [5-8].

Although preventive interventions for children of parents with mental illness or addiction are scarce, some interventions have been developed for this vulnerable group [9,10]. Most of these interventions are intended to be performed face-to-face and only a few are developed for online use. However, given the nature of the target group (eg, including minors, being at risk) and the problem being addressed, a face-to-face intervention is associated with numerous challenges regarding recruitment and inclusion. National Dutch data show that with current face-to-face interventions, the target population was not reached sufficiently and information did not find its way to the population [11]. The youth that were reached valued their anonymity and privacy, which makes face-to-face interventions less appealing. Therefore, online interventions seem to be a worthy alternative.

In the Netherlands, 864,000 parents meet the diagnostic criteria of a mental illness or addiction [12,13]. An online preventive course (Kopstoring) for children of these parents was developed. Kopstoring is one of the few online interventions for children of parents with mental illness or addiction disorders. The course is based on evidence-based theories and a face-to-face course developed for the same population. The Dutch Kopstoring course is designed for adolescents from ages 16 to 25 years. A pilot study assessing the effects of the course showed Kopstoring to be effective in improving participants' coping and mastery mechanisms [14].

The objective of this study is to gain knowledge about expectations, experiences, and perspectives of participants and providers of the online Kopstoring course. The research questions were how was the process of the delivery of the online Kopstoring course perceived by Kopstoring participants and providers and what were their expectations and experiences with this course?

Methods

A process evaluation was performed to evaluate the online delivery of Kopstoring and the experiences and perspectives of Kopstoring participants and Kopstoring providers. Interviews were performed with both groups. This section first describes the Kopstoring course and the accompanying randomized controlled trial (RCT; trial registration: NTR1982) [15] and subsequently the methods used in this process evaluation.

Intervention

The Kopstoring course aimed to prevent behavioral and psychological problems in offspring at risk and was offered to adolescents from ages 16 to 25 years. The Kopstoring course was an online 8-week group course with supervision by 2 trained psychologists or social workers from a participating mental health institution in the Netherlands. Every week a different theme was discussed and participants were expected to prepare for the weekly meetings by doing homework. The course had a preventive nature; therefore, adolescents were screened to ensure that they were not diagnosed with an illness as classified by the *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (*DSM-IV*) diagnoses. Screening was executed by the mental health institutions. In addition, participants needed to have access to a computer with an Internet connection and be able to participate weekly.

Alongside the process evaluation described in this paper, a RCT was conducted. The aim of the RCT was to examine the effectiveness and cost-effectiveness of the Kopstoring course. Participants were randomly allocated either to immediate enrollment in the Kopstoring course (intervention group) or enrollment after a 6-month waiting list (control group). Because the course was completely digitalized, the recruitment was done mainly, but not exclusively, through online recruitment, including banners, Facebook advertisements, links to the website, etc. In addition, articles were published in national and regional magazines and newspapers, school visits were performed, and an interview was broadcasted on a radio station.

Sample

In this process evaluation, 2 groups were included: participants (n=13) and providers (n=4) of the Kopstoring course.

The Kopstoring participants were selected from both the intervention and the control group of the RCT. Participants received an email in which they were invited to be interviewed.

To select participants for the interviews from the trial pool, a maximum variation strategy was used to gather information from a sample with as much variation as possible to collect as many different perspectives [16]. This was done by looking at several characteristics (eg, trial arm, age, sex, online and written consent for the trial, dropout). See Table 1 for the characteristics of the Kopstoring participants who were interviewed. The interview sample is fairly comparable to the RCT sample.

Table 1. Characteristics of interview participants versus trial participants.

Characteristics	Interview participants, n (%) n=13	Trial sample, n (%) n=104
Age [a] (years)		
16-17	3 (23)	20 (19.2)
>18	10 (77)	84 (80.7)
Sex		
Female	12 (92)	93 (89.4)
Male	1 (8)	11 (10.6)
Treatment group		
Intervention	8 (62)	55 (52.9)
Waiting list control	5 (38)	49 (47.1)
Adherence to Kopstoring course [b]		
Completed	11 (85)	97 (93.3)
Started but did not finish	2 (15)	7 (6.7)

[a] Age at time of registration for the course.

[b] To this point, data were checked up until 6 months after registration due to the pending follow-up assessments.

At the start of the project, 9 Dutch mental health institutions participated and each institution trained 2 professionals for the provision of the Kopstoring course. All planned Kopstoring courses provided during this study were provided by 7 professionals from 4 different mental health institutions. These providers were invited to participate in interviews. In total, 4 providers agreed to participate in an interview.

The providers of the Kopstoring course were all female and approximately 30 years of age; all had a Master’s degree and had experience working in this field for 5 to 7 years.

Data Collection

Data were collected through semistructured individual interviews with a list of topics to be discussed (Textbox 1). Interviews were held between November 2014 and February 2015. The topic list was made by the research team in collaboration with the national coordinator of Kopstoring and the team of course providers.

Textbox 1. Interview topic list for participants.

Contextual information

- Situation analyses
- Context analyses

Website Kopstoring

- Views about website
- Information delivery
- Logistics website

Effects of the course

- Views about the course
- Anticipated effects
- Experienced effects before, during, and after the course
- Barriers and success factors for completing the course

Process and content-related aspects of the course

- Components (themes) of the course
- Tailor-made health care
- Technical aspects of delivery online

Research

- Understanding study aspect
- Motivation
- Experience

Anonymity was very important for offspring of parents with mental illness or addiction problems. Therefore, interviews were conducted over the phone. Participants decided the time of the interview so they could be sure they were able to talk freely. Interviews with the providers were also conducted over the phone, but due to time constraints not because of anonymity. Interviews were held in Dutch. Textbox 2 displays the topic list for the providers.

Analysis

The interviews were audiotaped and transcribed and identifiable information was removed to ensure anonymity. The interviews were analyzed by using inductive qualitative content analysis, specifically conventional content analyses [17]. This method helped provide an in-depth understanding about underlying perspectives and qualitative methods are inductive and reflexive and it allowed the use of quotes [18]. As a first step, the interviews were read by 2 researchers separately to identify emerging themes and subthemes and then labels were attached to the parts related to these themes. Secondly, new themes were added to existing themes and labeled accordingly. After the 2 researchers reached consensus, the interview data were clustered into themes and subthemes. Finally, citations of the interviewees were identified per theme and visualized in a data matrix. After approximately 13 interviews, no new information emerged from the interviews with the participants.

Results

Five main themes emerged from the interviews: (1) background, (2) the requirements for the intervention, (3) experience with the intervention, (4) technical aspects, and (5) research aspects. In this section, each theme and its subthemes will be discussed from the participants' and the providers' perspectives.

Background

For the participants, the background mainly related to the motivation and reason for participation, the route to registration, and expectations of the online course. The providers' background related to their experiences with the provision of similar face-to-face courses and online interventions.

Textbox 2. Interview topic list for providers.

Contextual information

- Mental health institution
- Personal background
- General impression Kopstoring
- Financial situation and implications (Mental health institution)

Website Kopstoring

- Views about website
- Information delivery
- Logistics website

Provision, process, and content of the course

- Views about the content of course
- Advantages and disadvantages of online delivery
- Experiences with provision of Kopstoring
- Process from registration to allocation in group
- Barriers and success factors for provision of the course
- Technical aspects of delivery online

Research

- Motivation to participate
- Experience with the study

Participants

Participants from the Kopstoring course generally had 2 routes to arrive at the point of registration. Analyses showed that it was either a slow, lingering process in which the person already had the intention to change the situation for some time and was looking for a suitable way to address their needs or there was an acute situation which forced them to seek help right away. The following is an example of an acute situation that led to an immediate online registration:

> *There was a real occasion leading to why I registered. It was September last year and my mother had a psychosis...and she attacked me that night.* [Participant 12]

A respondent for whom the situation was ongoing long before registration explained:

> *I had a difficult time dealing with the situation and with the fact that my brother was placed into care (out of house placement). Well, I really could not handle it well, so they advised me to register for the course.* [Participant 4]

There was no difference in results reported by participants who registered under pressure of an acute event or those who took their time to register for the course.

For both situations, there appeared to be several facilitators, for example, a psychologist, school mental health worker, or a family member pointing out the online course or participants who found the course through an Internet search. Despite the different problems and family situations of participants (eg, one person had an addicted mom, another had a mentally ill father, and both parents were mentally ill for a third person), the consequences, questions, and problems they were confronted with were very similar.

Motivations to participate could be divided into 4 categories: (1) sharing experiences with persons in the same situation, (2) learning how to cope with ill parents, (3) learning how to cope with their own problems, and (4) learning about mental illness or addiction:

> *I hope it will be comforting for me to talk about my experiences with peers who went through the same experience.* [Participant 11]

Most of the time, problems were not discussed with family members and friends. This explained their need for sharing with others who had been through the same experience. One participant explained this as:

> *In a certain way, it provoked a sense of relief learning that other people were actually going through the exact same experience.* [Participant 4]

All participants had easy access to the website and experienced no problems with the registration process. During this process, all prospective participants were asked about their expectations of the online course. The answers were concise for the most part and participants had clear expectations regarding the content and the anticipated effect of the course.

When asked about their goals, participants reported they expected to learn and understand more about their parent's illness or addiction, to learn how to cope with the illness or addiction, and to learn how to improve the situation at home and decrease problems themselves.

One respondent explained that she hoped to find out how the situation got this "extreme" and to learn to deal with her mom so that they could improve their relationship and the situation. In accordance, other participants explained:

> *I expect this to be very helpful, mainly that I have more understanding about the cause of the symptoms and how to deal better with my mother. I hope to learn ways and also to detach myself from my mother.* [Participant 10]

Prospective participants, aged 16 to 25 years, had well-defined expectations of the online delivery and effects of the Kopstoring course.

Providers

Providers' general impression of the online delivery of Kopstoring was positive:

> *It is just a very good program that does not require any change. That is, of course, very important.* [Provider 1]

In some cases, the Kopstoring manual was considered theoretically so well written that providers used the same manual for face-to-face groups for children of mentally ill or addicted parents:

> *Yes, I find that very good [the course manual/protocol]. I even use it as the manual in the face-to-face Kop-groups. This is because I consider it to be a very pleasant way how subjects are being discussed, which themes will be covered, like cognitive behavior therapy.* [Provider 3]

Requirements

There are some requirements when providing online interventions. Firstly, participants and providers needed a computer connected to the Internet, log-in codes, and some privacy. For the providers, a budget was necessary to provide the Kopstoring course. Some barriers were encountered at the organizational level and regarding the financial structure.

Participants

In general, there were no barriers encountered to meet the requirements; however, one of the participants mentioned that when moving house she did not have access to the Internet, which made it impossible to log on to that session.

A second person explained that it was not always easy to find enough privacy in the house because there were always people around who did not know he was participating in the course:

> *The only thing that was difficult was finding a place to separate myself from others, and just having a moment for myself. That was difficult.* [Participant 12]

Providers

Providers needed support from their managers within the mental health institution and adequate finances to provide the online course because online mental health interventions are not paid from public funds. In addition, providers of online anonymous interventions face the situation that costs will not be reimbursed by insurers due to the fact they cannot provide social security numbers or other personalized details. It is up to the management of a mental health institution to decide whether or not it is feasible to add Kopstoring in their portfolio. In addition, the financial situation and structure of mental health care divisions for minors (up to age 18 years) changed during the course of the RCT. In short, municipalities became responsible for the policy and execution of budgeting of prevention interventions in youth mental health care. This shift had tremendous consequences for the delivery of mental health care interventions for youth up to age 18 years. In some institutions, prevention and youth departments were declared redundant and, consequently, the institutions withdrew their consent to provide the Kopstoring course. There were many problems encountered with finding funding to provide the Kopstoring course. One provider explained:

> *Health insurers are not paying for delivery of Kopstoring because it is provided anonymously and a health insurer only wants to reimburse when they have all details from the client. So that means you have to provide them with a health insurer registration number and social security number, everything, and we do not ask these details when providing Kopstoring because we want it to be anonymous. So the only remaining source is the municipality and...naturally the municipality actually only wants to pay for inhabitants of that municipality.* [Provider 1]

Initially, the costs of the courses were reimbursed by additional funding obtained by the research team. This meant that when the study period ended, reimbursement of the courses also came to an end. One provider explained that once the research team did not fund provision of the courses anymore, their mental health institution stopped providing the Kopstoring course.

Experiences with the Course

Participants

Participants described many different effects on their daily life and their problems. The first and most emphasized effect of the course was peer contact. Speaking with youth in the same situation made participants feel less alone, relieved, and less guilty in some cases. The recognition of situations, problems, and decisions became something they could share with peers:

> *My friends did not understand me. I have tried to explain, but then they would just say "awhhh it will be fine" and that was so nice to...with peers who might have a slightly different situation maybe, but that they also felt lonely and that you...share the same and have compassion for one another.* [Participant 5]

A second component that was considered very effective was the psychoeducative part of the course. Learning about the illness or addiction of the parent gave insight into the behavior of the parent:

> *I noticed that I experienced more peace with the fact that she has a drinking problem. It is the way it is and that will not change anymore.* [Participant 3]

Furthermore, the participants learned tips and tricks on how to cope with the behavior and problem of the parent, which led to accepting the parent's problem and more peace in the family in some cases:

> *Yes, I argued with my father pretty often if he had something going on. They said I should actually try to reduce these moments and I am able to do that now.* [Participant 1]

Participants also reflected on the content of the Kopstoring course and all themes were deemed important for the course to be effective. Almost all interviewees pointed out that the "rate your week" component, which kicked off every session, was extremely valuable to them. "Rate your week" was a simple but effective way to share experiences about the past week and a platform for questions and peer contact. A respondent explained the working mechanism of "rate your week":

> *Rate your week was very interesting for me. For your self-reflection on whether or not the week went well and that you were able to look back later to see how it was in the beginning of the course and how am I doing now? I really liked that; what are the positive things that make you also feel very positive instead of focusing on the negative things.* [Participant 9]

However, one of the sessions in which the educational component was key was mentioned to be a bit repetitive.

Experiences with the online program also translated into barriers. Barriers that existed were lack of time to discuss the homework assignments, some participants mentioned the course focused too much on the younger participants (students), and a couple of participants mentioned the homework was complicated.

Facilitating factors to adhere to the intervention and study were also mentioned. The online delivery of the intervention was mentioned as a major facilitating factor to start and finish the course. Firstly, online delivery was found to be convenient and ideal for participation in a safe and self-chosen environment. Secondly, online delivery protected the anonymity and privacy of the participant, which encouraged the participants to be more open:

> *Openness, yes...because it is online you do not have the feeling that everyone was looking at you. Then you can just write and maybe if you had to cry or so...no one was able to see that.* [Participant 8]

An often-mentioned stimulating factor was the attitude of the provider. Most of them were easy to access and always available (by email) to answer questions and monitor the participant:

> *I noticed that whenever the trainer tells us that she is still available to answer questions after the end of the course, or emails or these kind of things, that felt incredibly nice, that someone is still there who takes time, yeah where you can lean on. So that I consider to be very pleasant.* [Participant 6]

Only one participant mentioned that the attitude of the provider was not meeting her expectations. This participant stopped participating in the course after session 3 and was not included in the RCT.

Providers

All providers were satisfied with the content and agreed that all the important concepts were covered. The most important aspect was considered to be the online delivery, which ensured anonymity for the participant:

> *Within the Kopstoring course, they [participants] can, of course, tell their story very anonymously. Nobody knows that you participate in the group and what is bothering you. That is a huge advantage; that it becomes easily accessible for youngsters, but that they nevertheless can benefit and become more aware of what is going on and get answers to their questions.* [Provider 2]

For the content, the most important part was considered to be the exchange of experiences during "rate your week" at the beginning of each session.

All providers were asked several questions to check adherence to the protocol. They indicated they followed protocol except for one rule: the protocol described delivery of the course should be done by 2 professionals together. In practice, all providers delivered the course individually due to cost reductions. This was, however, not considered to be a barrier because providing the course for a group of up to 6 participants was highly manageable for one provider. There were some (technical) barriers experienced for the online delivery, but for providers the main barriers were experienced in the financial administrative field.

Positive factors were described as the feasibility of the online delivery and the possibility to deliver the course from home, the interactive group process, and the growing number of participants:

> *That every time again I am so surprised how close a group can become online and that as quick as in the first session they are already so open. And that is due to the anonymity that participants are just so open and what they think or experience...Yes, I think that this is very special and that stimulates me to provide the course over and over again and just getting back from them that they appreciate being heard.* [Provider 1]

Technical Aspects

The technical component was found to be extremely important by both the participants and providers of the Kopstoring course. Not only were the technical aspects (eg, the website or the chat box) considered positive factors, the same technical aspects were mentioned as barriers for participating in or providing the course. Almost every interviewee mentioned technical problems of some degree (from having a slow system to being thrown out of the chat box) and providers also mentioned these disturbances interrupting the courses. However, there were no major incidents mentioned that fully hindered participation in the long run.

Participants and Providers

The delivery of Kopstoring online was considered very positive; however, it seemed to also cause problems. Online delivery can be a double-edged sword; the convenience of the online aspect can be a pitfall at the same time.

Technical aspects were the design of the website and the chat box. Both were considered well designed and suitable for the target population. The website was described as complete, clear, colorful, and cheerful, which for participants is important. Providers shared this opinion about the website:

> *Just by clicking on the website, I consider it quite clear...I think it is convenient when registering for the course that the data when courses are starting is visible.* [Provider 2]
>
> *I considered it [chat box] very well done that different persons were indicated with different colors, so you were able to see who...to increase visibility.* [Participant 3]

Barriers

The technical problems described by both groups ranged from technical hiccups to some more prominent problems, such as not being able to log on to group sessions or being locked out by the system.

Minor technical problems involved a slow system, not being able to see when someone was typing, unable to see homework assignments on the screen, and double messaging occurring. These problems were mentioned, but were interpreted as minor problems and a consequence of an online working environment:

> *Sometimes it took like a minute or so before the text would be displayed or then it got stuck or we were removed from the chat box. Yeah every now and then we would struggle a bit.* [Participant 12]

Providers also described experiencing the same minor problems. The technical platform and responsibility related to the technical aspects were more numerous for the providers. They were responsible for all requirements to be met even before groups started online sessions.

Research Aspects

Participants were confronted with aspects such as a 6-month waiting list, randomization, extended follow-up, and questionnaires. It appeared that most participants understood there was a study linked to the Kopstoring course, but none of them could describe what the consequences were for them; regardless, patient information sheets were given to them by email, mail, and online:

> *I did know there were more groups where you could be allocated to, but I did not know that there was a chance you would have to wait half a year.* [Participant 12]

Participants who were allocated to the waiting list believed they had to wait because the group was full. This explained why most participants expressed no strong negative experiences toward the research components. In some cases, the waiting list was experienced as problematic, although most participants accepted the waiting period:

> *Yes, there was one group and I was hoping I could start right away, but unfortunately no. I had to wait half a year. That was really annoying. I needed help at that particular moment.* [Participant 8]

Most of the participants indicated altruism as the main reason for participating in the study, although others participated because “it is part of the course”:

> *I just hope that there are more young adults who get this opportunity to participate in this kind of course...that here is research, because yeah I feel that there is too little for Kopp? (children of mentally ill or addicted parents, for those groups).* [Participant 9]

Participants provided feedback on the length of the questionnaires and some technical problems related to not being able to open links or links expiring due to waiting too long to fill out the questionnaire. From the interviewed Kopstoring participants, 2 persons had incomplete data; when asked why, there was no specific reason, but they said they forgot. In addition, some of the respondents mentioned the phrasing of some of the questions. They disliked the questions because they were too focused on the younger participants (students) living with their parent(s).

During the interviews, a couple reasons for the poor response rates were mentioned: laziness or forgetfulness and problems with parental informed consent in case of a minor.

Participants were asked to give online consent and written consent sent by post. Only one participant did not send back the informed consent papers and, therefore, was not a participant in the RCT. She explained that she forgot to send the forms back, whereas the other participants had clear motivations for participation in the study. Some minors sent back their informed consent papers, but not those of their parents, stating they did not wish their parents to know let alone sign a consent paper for participation in the study. Some minors found a way to let their parents sign.

Providers

For providers, the research aspects were proper barriers. The back office and technical interface changed so much with the migration to the study environment that the coordinator was forced to invest time into teaching herself the new system. Providers also pointed out that the inclusion of participants in the study would have been smoother if the research team had listened to their advice concerning some of the requirements for study inclusion:

> *What I consider to be a real pity is that not enough weight is given to the advice of the professionals from the mental health institutions. ...Yeah, youngsters score incredibly high...if you are shocked by that as*

> *a researcher...It doesn't surprise or shock me anymore because I know that they score high, and I think it is very important to just reach out to these youngsters for participation in Kopstoring and motivate them to seek help.* [Provider 1]

One other barrier mentioned was the frustration they had to deal with when a person was allocated to the waiting list control group. Providers explained, “It is the way it is when doing research but it remained sad you had to disappoint a person who needs help right there and right then.”

Every 3 weeks, a new group was started. In addition, due to the allocation of participants to the waiting list, the number of people in the groups was reduced. Starting with smaller groups was considered a disadvantage because it was difficult to reach the full potential of the course.

Regardless of these barriers, providers were determined to reach the target number of participants for inclusion in the RCT. Their motivation was based on several stimulating factors. One provider explained that her opinion about the value of the product made her enthusiastic to provide the course and help with inclusion in the study:

> *I think you are enthusiastic if you see the value of research. That will lead to results and, surely, I hope the results are good. You also notice that participants have very positive experiences and so you feel you are working/providing a good product, so I think it matters a lot and the fact that it is studied, I only cheer for that.* [Provider 4]

Also, the interest and engagement of the researcher played a role in the delivery of the Kopstoring course and the willingness to help:

> *I enjoyed that you (researcher) were present at all meetings and gave an update on how the situation was and, yes, then we had an idea of what the situation was and that is what you are working towards.* [Provider 2]

Providers explained there were barriers; despite these, they were able to work within the study parameters.

Discussion

Principal Results

To our knowledge, this study is the first evaluation of the experience of providers and participants with an online-delivered prevention course for offspring at risk. Therefore, this study differentiates itself from existing international literature and provides new information. The few process evaluations performed to assess experiences with online programs focus on online treatment, programs for somatic diseases, or and/or an adult patient population. The findings of this study give insight into the experiences of participants and providers of an online prevention course called Kopstoring. It sheds some light on experience with as well as barriers and facilitating factors of online delivery. It elaborates on the expectations and experiences of both participants and providers. Analyses showed similar experiences for the 2 groups despite their different perspectives.

The main lesson learned from participants lies in their assessment of the course content and the barriers and facilitating factors for participating and adhering to an online course. The online aspect and anonymity proved to be important as well as their autonomy to decide to participate without interference from anyone else. In the Netherlands and some other countries, minors (participants younger than 18 years) need to provide the research team with written consent and their parent's consent for participating in a scientific study. This ignores the fact that minors can receive treatment (which is being assessed) of any kind from the age of 16 years without parental consent. Youth are considered capable of making an informed decision about treatment; however, for a scientific study we doubt their ability to make an independent and informed choice [19]. This subject was brought up even by participants older than 18 years who said that if they had to provide parental consent, they would probably not have participated. They explained that we cannot expect them to ask their parents for consent when they are the root of their problems. There is a strong need to rethink the policy concerning consent in the case of interventions for vulnerable populations and interventions with a high level of anonymity (mostly online interventions). This statement endorses the debate in the literature questioning when a minor should be considered capable to give informed consent and therefore protect his/her anonymity [20,21]. Despite the differences in consent procedure for minors, no substantial differences between minors and participants older than 18 years are reported in this study.

The lesson learned from the providers of the Kopstoring course lies partly in their professional assessment of the content of the course, but mainly in the experience with providing online courses and the barriers and facilitating factors to provide the Kopstoring course. Analyses showed that providers of online interventions in RCTs might feel ignored and may experience a gap between the research team and providers, even though the provider has many years of experience with providing online interventions in this target population. This implies that there is a need for closer collaboration with providers, and perhaps even with the target population, when designing such interventions and accompanying studies [22,23]. Collaborating with stakeholders could have led to other research questions, methods, and the use of other questionnaires more suitable for the target population.

The current situation in the Netherlands for children in need of mental health care is unsettling. This study could not have taken place in a more inconvenient time and political setting than it actually did. In the same period the RCT was running, political decisions forced youth mental health care out of the hands of mental health institutions and made it subsequently a part of the local municipalities. Even if the results of cost-effectiveness studies, such as the Kopstoring RCT, show positive results, online interventions may not be provided due to the complex financial structure and lack of responsible bodies to finance online interventions. This also shows that implementation and implementation research in the Netherlands, but very likely in other similar countries, is nearly impossible for these types of intervention.

Limitations of the Study

There are several factors that could be considered to influence the findings of this study. The first is the number of interviews performed. One can question whether the small number of providers interviewed is sufficient to provide a complete overview of the ongoing issues. However, we do believe that providers who cooperated gave a lot of information about the Kopstoring course and the delivery of the course. We remained with only 9 providers who provided more than one course and were totally informed about every research detail. Therefore, we believe 4 providers were a good reflection of the 9 remaining providers and the group appeared to be homogeneous.

For the participants, there is a different reason for the low response rate (42 people were invited, 17 responded, and 13 were eventually interviewed; response rate: 13/42, 31%) for participating in this study. The target population appears to be extremely difficult to reach. As shown in the analyses, they wish their anonymity to be respected and feel “safe” in an online environment and not face-to-face or on the phone. In addition, a feeling of shame and guilt regarding their problems blocks them from sharing their experiences with a researcher. Despite this, the majority of the participants were enthusiastic about the online prevention course and potential bias might occur with this. It is possible that youth with negative experiences with the course or research were not willing to be interviewed. Additionally, due to the sensitivity of this problem and the fact that the parents are involved, youth might find it difficult to speak about this with a third party (ie, might feel like “airing their dirty laundry” in public). However, for both participants and providers, repetition in the interviews showed a level of saturation.

A second limitation relates to the generalizability of the findings. It is noticeable that an overwhelming majority of Kopstoring participants, participating in the underlying RCT and this process evaluation, were female. This is probably not a good reflection of an open population, assuming there are an almost equal number of boys who have a mentally ill parent as there are girls. This leaves questions about generalizability unanswered. In addition, questions have been raised such as “are the findings useful in a similar online context, but with a different underlying intervention?” and “are the findings the same when comparing the online course to a similar face-to-face group?” It appeared that several factors added up; the online aspect, age, anonymity, and sensitive problems and anonymity lead to barriers doing research within this vulnerable group. The results of this study focus on youth with parents with mental illness or addiction problems. Despite these factors, some general elements can be identified that are useful in other online settings, such as the aspect of anonymity, consent, and practical issues.

Conclusions

Online support for offspring of parents with mental illness or addiction problems is considered effective by the participants. There are not many RCTs performed to assess the effectiveness and cost-effectiveness of online prevention programs in the field of mental health care [24]. Consequently, there are not many process evaluations of these online prevention programs performed. This hampers comparison between online programs and process of delivery and expectations. In addition, a face-to-face group is set up differently in structure and has fewer participants; therefore, it is difficult to use it in comparison to an online program. In this respect, this study is unique and sheds some light on experiences and barriers for online provision of a prevention course in the field of mental health care.

The barriers for online provision of this health intervention are minimal, but the ones that exist lie in the technical sphere. Barriers for online research are multiple and touch on different aspects, such as informed consent, anonymity, lack of time, or just lack of interest. The findings of this study may explain partly why there are substantial dropout rates when delivering online interventions. The experiences of participants and providers of the Kopstoring course give valuable insights into the process of the online provision and study of Kopstoring.

Acknowledgments

The authors would like to thank all participants from the Kopstoring RCT and the ones who were willing to be interviewed for the process evaluation. In addition, the authors would like to thank Karlijn Arntz who has done an outstanding job as coordinator of the participating mental health institutions. We would also like to thank the participating mental health insitutions: Eleos, Dimence, Mindfit, GGNet, Vincent van Gogh Instelling & Riagg Zuid, and Ypse Dixet for providing the Kopstoring course. This study was an addition to the Kopstoring RCT which is financed by ZonMw (The Netherlands Organisation for Health Research and Development grant number: 200210002) and provision of the Kopstoring course was funded by the Dutch Innovatiefonds Zorgverzekeraars.

Conflicts of Interest

None declared.

References

1. Ustün T, Ayuso-Mateos J, Chatterji S, Mathers C, Murray CJ. Global burden of depressive disorders in the year 2000. Br J Psychiatry 2004 May;184:386-392 [FREE Full text] [Medline: 15123501]
2. Maurin J, Boyd C. Burden of mental illness on the family: a critical review. Arch Psychiatr Nurs 1990;4(2):99-107. [Medline: 2192692]

3. Bijl R, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 1998;33(12):587-595. [Medline: 9857791]
4. Bassani DG, Padoin CV, Philipp D, Veldhuizen S. Estimating the number of children exposed to parental psychiatric disorders through a national health survey. Child Adolesc Psychiatry Ment Health 2009;3(1):6. [Medline: 19228427]
5. Beardslee WR, Gladstone TR, O'Connor EE. Transmission and prevention of mood disorders among children of affectively ill parents: a review. J Am Acad Child Adolesc Psychiatry 2011 Nov;50(11):1098-1109. [doi: 10.1016/j.jaac.2011.07.020] [Medline: 22023998]
6. Weissman M, Wickramaratne P, Nomura Y, Warner V, Pilowsky D, Verdeli H. Offspring of depressed parents 20 years later. Am J Psychiatry 2006;163(6):20-28. [Medline: 16741200]
7. Weissman M, Fendrich M, Warner V, Wickramaratne P. Incidence of psychiatric disorder in offspring at high and low risk for depression. J Am Acad Child Adolesc Psychiatry 1992 Jul;31(4):640-648. [doi: 10.1097/00004583-199207000-00010] [Medline: 1644726]
8. Downey GJ, Coyne JC. Children of depressed parents: an integrative review. Psychol Bull 1990;108(1):50-76. [Medline: 2200073]
9. Siegenthaler E, Munder T, Egger M. Effect of preventive interventions in mentally ill parents on the mental health of the offspring: systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 2012 Jan;51(1):8-17.e8. [doi: 10.1016/j.jaac.2011.10.018] [Medline: 22176935]
10. Fraser C, James E, Anderson K, Lloyd DF, Judd F. Intervention programs for children of parents with a mental illness: a critical review. International Journal of Mental Health Promotion 2006 Feb;8(1):9-20. [doi: 10.1080/14623730.2006.9721897]
11. De Jonge M, Ruiter M. Aard en Omvang GGZ- en Verslavingspreventie 2008- Ontwikkelingen in de inhoud, organisatie en omvang van de GGZ- en verslavingspreventie. Netherlands: Trimbos Instituut; 2010.
12. van Dorsselaer S, Zeijl E, van den Eeckhout S, ter Bogt T, Vollebergh W. HBSC 2005: Gezondheid en welzijn van jongeren in Nederland. Utrecht: Trimbos Instituut; 2007.
13. Bool M, van der Zanden R, Smit F. Factsheet preventie: Kinderen van ouders met psychische problemen (Prevention factsheet: Children with mentally ill parents), T Instituut, Editor. Utrecht: Trimbos Insituut; 2007.
14. van Veen C, van der Zanden R. Procesevaluatie Kopstoring; Een onderzoek naar de preventieve online groepscursus voor jongeren van 16-25 jaar. Utrecht: Trimbos Instituut; 2007.
15. Woolderink M, Smit F, van der Zanden R, Beecham J, Knapp M, Paulus A, et al. Design of an internet-based health economic evaluation of a preventive group-intervention for children of parents with mental illness or substance use disorders. BMC Public Health 2010;10:470 [FREE Full text] [doi: 10.1186/1471-2458-10-470] [Medline: 20698962]
16. Coyne IT. Sampling in qualitative research. Purposeful and theoretical sampling; merging or clear boundaries? J Adv Nurs 1997 Sep;26(3):623-630. [Medline: 9378886]
17. Hsieh HS, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9):1277-1288. [Medline: 16204405]
18. Coast J, McDonald R, Baker R. Issues arising from the use of qualitative methods in health economics. J Health Serv Res Policy 2004 Jul;9(3):171-176. [doi: 10.1258/1355819041403286] [Medline: 15272976]
19. Dickens BM, Cook RJ. Adolescents and consent to treatment. International Journal of Gynecology and Obstetrics 2005;89:5. [Medline: 15847892]
20. Alderson P. Competent children? Minors' consent to health care treatment and research. Soc Sci Med 2007 Dec;65(11):2272-2283. [doi: 10.1016/j.socscimed.2007.08.005] [Medline: 17854964]
21. Woolderink M, van Asselt AD, van Schayck CP, van Wijmen FB. E-research: problems with anonymity and consent. Ned Tijdschr Geneeskd 2013;157(51):A6828. [Medline: 24345359]
22. Abma T, Nierse CJ, Widdershoven GA. Patients as partners in responsive research: methodological notions for collaborations in mixed research teams. Qual Health Res 2009;19(3):401-415. [Medline: 19224882]
23. Bindels J, Baur V, Cox K, Heijing ST. Older people as co-researchers: a collaborative journey. Ageing & Society 2014;34(06):951-973.
24. Calear AH, Christensen H. Review of internet-based prevention and treatment programs for anxiety and depression in children and adolescents. Med J Aust 2010;192(11 Suppl):S12-S14. [Medline: 20528700]

Abbreviations

RCT: randomized controlled trial

Edited by P Bamidis; submitted 13.06.15; peer-reviewed by P Sarafis, P Domeyer; comments to author 12.08.15; revised version received 14.09.15; accepted 28.09.15; published 02.12.15.

Please cite as:
Woolderink M, Bindels JAPM, Evers SMAA, Paulus ATG, van Asselt ADI, van Schayck OCP
An Online Health Prevention Intervention for Youth with Addicted or Mentally Ill Parents: Experiences and Perspectives of Participants and Providers from a Randomized Controlled Trial
J Med Internet Res 2015;17(12):e274
URL: http://www.jmir.org/2015/12/e274/
doi:10.2196/jmir.4817
PMID:26633244

Original Paper

Feasibility and Quit Rates of the Tobacco Status Project: A Facebook Smoking Cessation Intervention for Young Adults

Danielle E Ramo[1,2,3], PhD; Johannes Thrul[3], PhD; Kathryn Chavez[1], BA; Kevin L Delucchi[1], PhD; Judith J Prochaska[4], MPH, PhD

[1]Department of Psychiatry, University of California, San Francisco, San Francisco, CA, United States

[2]UCSF Helen Diller Family Comprehensive Cancer Center Tobacco Control Program, University of California, San Francisco, San Francisco, CA, United States

[3]Center for Tobacco Control Research and Education, University of California, San Francisco, San Francisco, CA, United States

[4]Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, CA, United States

Corresponding Author:
Danielle E Ramo, PhD
Department of Psychiatry
University of California, San Francisco
401 Parnassus Avenue
Box TRC 0984
San Francisco, CA, 94143
United States
Phone: 1 4154767695
Fax: 1 4154767053
Email: danielle.ramo@ucsf.edu

Abstract

Background: Young adult smokers are a challenging group to engage in smoking cessation interventions. With wide reach and engagement among users, Facebook offers opportunity to engage young people in socially supportive communities for quitting smoking and sustaining abstinence.

Objective: We developed and tested initial efficacy, engagement, and acceptability of the Tobacco Status Project, a smoking cessation intervention for young adults delivered within Facebook.

Methods: The intervention was based on the US Public Health Service Clinical Practice Guidelines and the Transtheoretical Model and enrolled participants into study-run 3-month secret Facebook groups matched on readiness to quit smoking. Cigarette smokers (N=79) aged 18-25, who used Facebook on most days, were recruited via Facebook. All participants received the intervention and were randomized to one of three monetary incentive groups tied to engagement (commenting in groups). Assessments were completed at baseline, 3-, 6-, and 12-months follow-up. Analyses examined retention, smoking outcomes over 12 months (7-day point prevalence abstinence, ≥50% reduction in cigarettes smoked, quit attempts and strategies used, readiness to quit), engagement, and satisfaction with the intervention.

Results: Retention was 82% (65/79) at 6 months and 72% (57/79) at 12 months. From baseline to 12-months follow-up, there was a significant increase in the proportion prepared to quit (10/79, 13%; 36/79, 46%, $P<.001$). Over a third (28/79, 35%) reduced their cigarette consumption by 50% or greater, and 66% (52/79) made at least one 24-hour quit attempt during the study. In an intent-to-treat analysis, 13% (10/79) self-reported 7-day abstinence (6/79, 8% verified biochemically) at 12-months follow-up. In their quit attempts, 11% (9/79) used a nicotine replacement therapy approved by the Food and Drug Administration, while 18% (14/79) used an electronic nicotine delivery system to quit (eg, electronic cigarette). A majority (48/79, 61%) commented on at least one Facebook post, with more commenting among those with biochemically verified abstinence at 3 months ($P=.036$) and those randomized to receive a personal monetary incentive ($P=.015$). Over a third of participants (28/79, 35%) reported reading most or all of the Facebook posts. Highest acceptability ratings of the intervention were for post ease (57/79, 72%) and thinking about what they read (52/79, 66%); 71% (56/79) recommended the program to others. Only 5 participants attended the optional cognitive-behavioral counseling sessions, though their attendance was high (6/7 sessions overall) and the sessions were rated as easy to understand, useful, and helpful (all 90-100% agreed).

Conclusions: A Facebook quit smoking intervention is attractive and feasible to deliver, and early efficacy data are encouraging. However, the 1.5-fold greater use of electronic cigarettes over nicotine replacement products for quitting is concerning.

(J Med Internet Res 2015;17(12):e291) doi:10.2196/jmir.5209

KEYWORDS

smoking cessation; Facebook; social media; young adults

Introduction

Although the prevalence of cigarette smoking has declined among US adults since 1983, the smoking prevalence among young adults aged 18-25 years has remained stable, with past month cigarette use rates as high as 31% in 2013 [1]. Compared to other age groups, young adults are less likely to use behavioral or pharmacotherapy interventions for smoking cessation [2], and studies of tobacco use have reported great challenges in recruiting young adults [3,4]. Reaching the US public health goal to reduce smoking prevalence to no higher than 12% by 2050 will require novel intervention approaches, enhancement of the effectiveness of existing treatments, and maximized reach and utilization of both.

Web-based social networks may serve as a solution to the problems of reach and engagement in smoking cessation interventions for young adults. Social networks, including those developed through online social media, play a role in onset and perpetuation of smoking behavior [5,6]. Patterns of social interactions through online smoking cessation forums have characterized ways that social networks can influence smoking cessation through social support, including by offering encouragement and emotional support, stories, congratulations, “thank you” messages, giving practical advice and tips, and discussing nicotine replacement therapy [7,8]. With wide reach and engagement among users, social media tools offer phenomenal opportunity to use social interactions to engage young people in behavior change interventions and to foster socially supportive communities for quitting smoking and sustaining abstinence.

With use continuing to increase annually, Facebook remains the most widely used social media tool and the second most popular website in the United States [9] with over 156 million US users as of January 2015 [10]. With 87% of US online young adults having a Facebook account and 70% of those accessing it daily [11], there is promise to use this platform to deliver public health intervention programs to young people.

Previous evaluations using Facebook to change health risk behavior have shown feasibility as measured by participant engagement and satisfaction [12-20]. However, trials examining social media interventions have shown limited or no effects on health behavior change (eg, physical activity) [21]. As applied to smoking, the BIO smoking cessation campaign for young adults in Canada, incorporating a website, smartphone app, and Facebook features, resulted in greater 7-day and 30-day reported quit rates than referral to a smokers’ helpline at 3-month follow-up [22]. Research is needed to determine whether Facebook alone can be used as an intervention tool for smoking cessation, whether abstinence can be biochemically verified, and whether abstinence rates can be maintained over 1 year.

In a mixed-methods study, we previously [23] examined 570 young adults’ receptivity to using Facebook to quit smoking, and 31% reported they would want to get help to quit smoking using Facebook. Interest in using Facebook to quit was greater among those more motivated to quit, who had made a quit attempt in the past year, and had previously used the Internet for assistance with a quit attempt. In qualitative interviews, social support and convenience were identified as strengths of a Facebook intervention, while privacy was the main issue of concern. It was determined that an intervention delivered through Facebook should be tailored to readiness to quit smoking and should deliver evidence-based content to groups of smokers who can support one another, while maintaining privacy from larger Facebook social networks.

While retention in social media intervention studies is promising, engagement may not be high enough to promote and sustain behavior change. Given that engagement in Web interventions influences efficacy [24], and social media is fraught with some of the same engagement concerns as online interventions with large drop-offs in participation among users [25], engagement is of utmost importance in designing this type of intervention. The use of monetary incentives is successful in recruiting participants to randomized clinical trials [26] and in yielding short-term abstinence to substance use [27]; however, the use of incentives to increase engagement in social media behavior change intervention has yet to be evaluated. What type of incentive intervention works best with young adults in a Facebook intervention is an empirical question. Monetary incentives have been effective in other settings, but it is also possible that donation would be a motivator for engagement in a health behavior intervention, especially for those with high amounts of trait altruism, an intrinsic motivator [28].

Herein, we describe feasibility of the Tobacco Status Project, a Facebook smoking cessation intervention for young adults in the United States. The intervention was tailored to readiness to quit smoking; therefore, participants did not have to want to quit smoking to participate. Goals of this trial were to examine feasibility and initial efficacy of the intervention, including participant characteristics, retention, intervention characteristics (eg, number of groups formed, size of groups), smoking outcomes over 12 months (7-day point prevalence abstinence, ≥50% reduction in cigarettes smoked, quit attempts and strategies used, readiness to quit), engagement, and satisfaction with the intervention. To inform best practices for engagement in a larger trial, we also compared incentive conditions contingent upon daily commenting to the study. Participants could either receive a personal incentive or an incentive for a charity donation in the participant’s name. Last, we tested moderating effects by trait altruism.

Methods

Participants and Recruitment

Participants were aged 18-25, English literate, and had smoked ≥100 cigarettes in their lifetime, currently smoked at least 3 days per week (consistent with the definition of smoking used in the National Health Interview Survey) [29], and used Facebook at least 4 days per week. Participants had to have access to technology that could take and send an online picture for verification of tobacco abstinence with a study mailed saliva cotinine test.

Recruitment efforts included a paid Facebook ad campaign conducted between June and August 2013 with details reported previously [30]. Advertisements invited participants to a secure, confidential online survey to evaluate eligibility and, for those eligible, informed consent to participate in the intervention. Online consent questions were used to confirm understanding of study procedures. Consented participants were asked to send proof of identity either by emailing a copy of a photo ID with birth date or by "friending" the study on Facebook to determine age.

Study Design

After completing an online baseline assessment, all participants were assigned to "secret" Facebook groups tailored to their stage of change for quitting smoking (preceontemplation, contemplation, preparation) [31,32]. As per Facebook privacy options for groups, only study administrators and group members knew the existence of "secret" groups, members of the group, or any member activity. This was distinct from a "public" group, in which anyone could join and actions were public, or a "closed" group, in which anyone could ask to join or be invited and existence/membership was not private. More information about the differences between group types on Facebook is described in Facebook's Help Center [33]. Secret groups were chosen based on mixed-methods formative work with young adults suggesting that privacy was an important consideration that would likely affect participation in a Facebook smoking cessation intervention [23]. A stage-matched group began 2 weeks after the first participant was assigned so that no group member was kept waiting longer than 2 weeks to start the intervention; thus, group sizes varied. The intervention was delivered daily for 90 days. During the intervention, participants' stage of change was assessed monthly and if they had advanced, they were invited to join a later stage group. As this was a feasibility evaluation, participants could join the later stage group, remain in their original group, or add a second group and participate in two. We assessed program use and acceptance at intervention end (3-month follow-up), and smoking outcomes at 3, 6, and 12 months. Participants were contacted by email and Facebook to complete online follow-up assessments and compensated for each assessment with US $20 gift cards to their choice of national retailer (eg, Amazon, Best Buy) and an additional US $20 gift card if all four assessments were completed for a total possible compensation of US $100.

Intervention Description

Secret Groups

All participants received access to a secret Facebook group tailored to their stage of change: Precontemplation: "Not Ready to Quit"; Contemplation: "Thinking About Quitting"; or Preparation: "Getting Ready to Quit." Participants were invited to participate in the secret group on Facebook. All groups received daily Facebook postings for 90 days tailored to their stage of change and consistent with US Clinical Practice Guidelines for Treating Tobacco Use and Dependence [34]. Those in the "Not Ready to Quit" groups received messages incorporating the 5Rs (relevance, risks, rewards, roadblocks, repetition) [35,36]; core Motivational Interviewing techniques of expressing empathy, rolling with resistance, supporting self-efficacy, and developing discrepancy [35,37-41]; and Transtheoretical Model strategies of increasing the pros of quitting, raising consciousness about quitting smoking, and environmental opportunities to quit smoking (eg, clean indoor air laws) [42]. "Thinking About Quitting" group posts emphasized decreasing the cons of smoking, and environmental re-evaluation (identifying negative effects of smoking on others and positive effects of change). "Getting ready to quit" posts focused on self-liberation (eg, making a commitment to quit), stimulus control (eg, removing smoking paraphernalia from the home), and counter conditioning (eg, engaging in alternative behaviors). The posts used a mix of imagery, text, and Facebook poll formats.

Ask the Doctor Sessions

In all groups, regardless of stage of change, weekly interactive sessions were conducted with the first author, during which participants could ask any questions related to smoking or quitting. All sessions included a post introducing the live, hour-long interactive session, and invited participants to share any questions or issues that they wanted to discuss regarding smoking or quitting. These sessions were not formally scripted to allow for participant-guided content. Responses to posts incorporated motivational interviewing in all groups, especially useful for younger smokers and those with low motivation to quit smoking [40,43] and cognitive behavioral counseling techniques for those ready to quit, recommended by US Clinical Practice Guidelines [34]. Initially, sessions were conducted at the same time each week based on participants' reports of when they were most likely to use Facebook. As the intervention progressed, session times varied. Sessions were scheduled and run through the secret groups.

Cognitive Behavioral Therapy Cessation Sessions

At any time during the intervention through the 12-month follow-up, participants could opt to participate in 7 sessions of group cognitive behavioral therapy (CBT) counseling delivered through Facebook chat with the first author. Interest was assessed through secret groups, and interested participants were scheduled for an initial session with the first author. The initial session was designed to help participants set a quit date and make a specific quit plan. Participants were then assigned to a CBT group based on availability of other interested participants (not necessarily in the same secret group) and asked to attend weekly sessions through Facebook chat. At the weekly time, the counselor invited all group members to a private chat session through Facebook's group chat feature. Content for sessions was scripted and adapted for social media delivery from a manual developed by Brown et al [40]. Sessions included text and images designed to be pasted into chat sessions by the counselor along with ad hoc responses to session-specific content. Topics in the six group sessions included (1) Preparing for Quitting; (2) Celebrating Cessation, Addressing Withdrawal & High-Risk Situations; (3) Getting Support and Asserting One’s Needs; (4) Managing Mood and Stress; (5) Living Healthy: Exercise, Food, and Substance Use; and (6) Maintaining Motivation, Graduation, and Looking Forward.

Although not directly available through the study, information was given about nicotine replacement therapy and medication for smoking cessation to all groups through posts, Ask the Doctor sessions, and CBT counseling sessions.

Intervention Engagement Incentives

Participants were able to interact with the intervention by “liking” or commenting on intervention posts, in Ask the Doctor sessions, or to make original posts or comment on other members’ posts. In a three-group design, we tested the utility of using incentives (none, personal, altruistic) to encourage intervention engagement. All participants were randomized to one of three incentive conditions: (1) Personal incentive: Participants were told that a US $50 gift card would be emailed if they commented on all 90 posts to their Facebook secret group by the end of the 90-day intervention period; (2) Altruistic incentive: A US $50 gift card to be donated to a charitable organization of their choice through the “justgive” website if they commented on all 90 posts; or (3) No incentive. This design was used to balance maximizing engagement (commenting on all 90 posts) with staff burden of counting all comments to Facebook group only once (at the end of the intervention period). Incentive condition was evaluated in relation to “likes” and comments, as well as 3-month abstinence. Incentive group was not hypothesized to have an effect on longer-term abstinence.

Upon completion of the intervention, data from secret groups (all likes and comments) were extracted from Facebook through the Facebook application programming interface.

Measures

All measures were administered online using Qualtrics. Qualtrics is a secure, online, Health Insurance Portability and Accountability Act (HIPAA) compliant survey software that transmits data to and from secure, firewalled data centers using Transport Layer Security encryption, the successor to Secure Sockets Layer encryption. Secure access is available to all faculty and staff at the University of California, San Francisco.

Demographics

At baseline, we assessed age, gender, race/ethnicity, household income, years of education, college enrollment, and employment.

Smoking

A Smoking Questionnaire [44] assessed average number of cigarettes smoked per day, and days smoking per week, total cigarettes smoked in the past week, age of initiation of smoking, years of smoking, and presence of past year quit attempt (y/n). Participants’ future smoking goals were assessed using one item with 7 response options, categorized into 3 categories: (1) no goal, (2) controlled or reduced smoking, and (3) abstinence [45]. Time to first cigarette upon waking (<30 min or >30 min) measured dependence [46]. The Smoking Stages of Change Questionnaire [31] was administered at baseline, 3, 6, and 12 months.

The primary outcome was tobacco abstinence, assessed and verified at 3, 6, and 12 months as the number of cigarettes smoked in the past week (even a puff; 7-day abstinence) or use of any form of tobacco. Consensus guidelines recommend use of 7-day point prevalence abstinence in cessation induction studies with smokers unmotivated to quit, who will be quitting at different time points in a trial (ie, cessation induction trials) [47]. Those reporting “no smoking in the past 7 days” were sent by mail a semiquantitative NicAlert saliva cotinine test kit with instructions to send two photos to research staff: one showing the participant putting saliva into a single collection tube and a second of the results. Cotinine levels of 30 ng/mL or less verified abstinence.

Secondary outcomes of interest were proportion of respondents reporting 50% or greater reduction in the number of cigarettes smoked in the past 7 days between baseline and 3 months, 6 months, and 12 months, and percentage of participants reporting a 24-hour smoking quit attempt during the study as well as the number of quit attempts reported. We also assessed strategies participants used to aid cessation attempts at 3, 6, and 12 months, including cessation medication or counseling (other than our intervention), or electronic nicotine delivery systems (ENDS). Any use over the 12-month study was determined (y/n) for each strategy.

Engagement

Engagement was examined by quantifying the number of “likes” and comments made during the intervention period. Engagement in the CBT counseling was assessed as the number of participants who opted for CBT sessions and the mean number of sessions they attended.

Program Acceptance/Likability

At intervention end (3 months), an 8-item questionnaire assessed participant reaction to the Facebook group (postings and “The Dr. Is In” sessions) and each CBT session, measured on a 4-point scale from “strongly disagree” to “strongly agree.” Proportions of those reporting “agree” or “strongly agree” were computed for each item.

Altruism

Altruism was measured using a hypothetical version of the Dictator Game, often used in economics experiments, in which participants are given an endowment and must choose the amount to allocate to themselves and how much to another person. In our version, participants were granted US $10 and asked to consider keeping it or sending any portion of it in $1 increments, to another anonymous person [48,49]. The amount sent was regarded as a measure of other-regarding or altruism (scored between 0 [least altruistic] and 10 [most altruistic]). See Multimedia Appendix 1 for the measure used.

Analyses

Descriptive statistics characterized the sample at baseline, evaluated the number of completed assessments at 3, 6, and 12 months, and characterized the intervention (number and size of groups, number of participants changing or leaving groups, number of participants opting to attend CBT sessions). Primary and secondary outcomes were assessed at 3, 6, and 12 months. Primary outcomes were 7-day point prevalence abstinence and biochemically verified abstinence. Secondary outcomes of interest were the proportion that made at least one 24-hour quit attempt, calculated for each time period, and for the entire 12-month follow-up period. Use of strategies to quit smoking between baseline and 12-month follow-up was summarized and predictors of strategy use were analyzed with logistic regression analyses. Wilcoxon signed rank tests compared past 7-day cigarettes smoked between baseline and each follow-up. Bowker's test for change in a repeated categorical measures evaluated stage of change transitions between baseline and 3 months, 3 and 6 months, and 6 and 12 months (3 models). To evaluate engagement, total number of "likes" and comments to Facebook groups over 3 months were tallied. Due to skew, the non-parametric Kruskal-Wallace analysis of variance (ANOVA) evaluated comments to the Facebook groups by baseline stage of change. Mann-Whitney U tests compared comments for those with biochemical-verification of abstinence at 3 months to those non-abstinent. Kruskal-Wallace tests evaluated the effect of incentive condition (personal, altruistic, no incentive) on comments to the Facebook group, in both the full sample and only those who made at least one comment to the group to address highly skewed data. Within the altruism incentive group, Kendall's tau test indexed trait altruism by comments. Kendall's tau test also indexed incentive condition by 3-month reported abstinence and biochemically verified abstinence. Usability and satisfaction with the intervention were evaluated by tallying the proportion of users answering "agree" or "disagree" to each item on the usability measure.

Results

Of the 586 respondents who met criteria to participate, 39.2% (230/586) signed online consent, 19.3% (113/586) verified identify online, and 13.5% (79/586) completed a baseline assessment and were assigned to a Facebook group (Figure 1).

Baseline Demographic and Smoking Characteristics

The 79 enrolled participants who completed a baseline assessment, on average, were 20.8 years old (SD 2.1) and primarily male (63/79, 80%) and non-Hispanic white (63/79, 80%). Seventeen percent (13/79) identified as sexual minority (lesbian, gay, or bisexual). About one quarter (22/79, 28%) had a household income >US $60,000. Average years of education was 12.4 (SD 2.0); 38% (30/79) were currently enrolled in school and 56% (44/79) were currently employed; and 43% (34/79) lived with their parents. Three quarters (59/79, 75%) were daily smokers, smoking 11.5 cigarettes per day on average (SD 8.3). Only a minority (8/79, 10%) of participants reported an abstinence smoking goal, 60% (47/79) reported a reduction goal, and 30% (24/79) reported no goal. More than half (41/79, 52%) of participants smoked within 30 minutes of waking. Participants smoked their first cigarette at age 14.2 years (SD 3.4), started smoking regularly at age 16.3 (SD 2.8), and had smoked for 2.7 years (SD 0.7) on average. Over half (45/79, 57%) had made at least one smoking quit attempt in the past year.

Retention

Follow-up completion was 76% (60/79) at 3 months, 82% (65/79) at 6 months, and 72% (57/79) at 12 months with no difference in 3-month retention by incentive condition ($\chi^2_{2,79}$=.31, P=.855).

Intervention Characteristics

Seven Facebook secret groups were created of varying sizes (two "Not Ready to Quit," three "Thinking About Quitting," and two "Getting Ready to Quit"; group size range 6-22). During the 90-day intervention period, 5% (4/79) of participants opted to change to a later stage group; 8% (6/79) of participants left their Facebook group completely at some point during the 3-month intervention period (3 precontemplation, 3 contemplation). Several participants (5/79, 6%) entered into CBT treatment, with one group made up of 3 participants, and 2 participants with individual treatments due to lack of other participants for groups. Participants attended 6 of 7 CBT chat sessions on average.

Primary Smoking Outcome

Assuming missing=missing (ie, complete case analysis), reported 7-day point prevalence abstinence was 12% (9/79), 22% (17/79), and 18% (14/79) at 3-, 6-, and 12-months follow-up (see Figure 2). Assuming those lost to follow-up were smoking (ie, missing=smoking), 7-day point prevalence abstinence was 9% (7/79), 18% (14/79), and 13% (10/79) at 3-, 6-, and 12-months follow-up. At each time point, cotinine-verified abstinence was confirmed with approximately half of those reporting abstinence.

Figure 1. Participant flow chart through a Facebook smoking cessation intervention. Those assessed for eligibility who were not counted as “excluded” left the survey too early to determine why they were ineligible (participants were randomized to incentive groups as follows: 24% Personal; 28% Altruistic; 48% None).

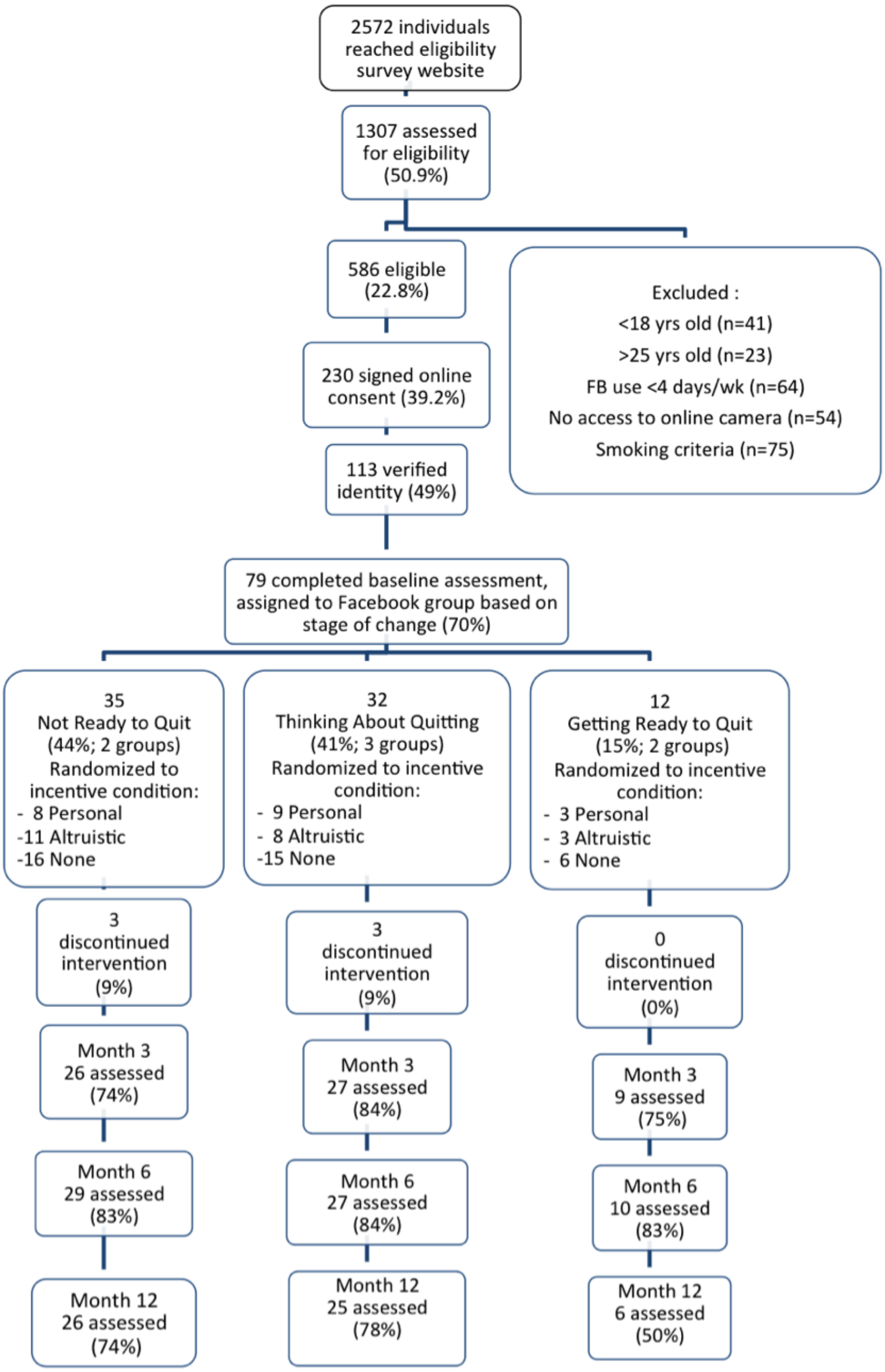
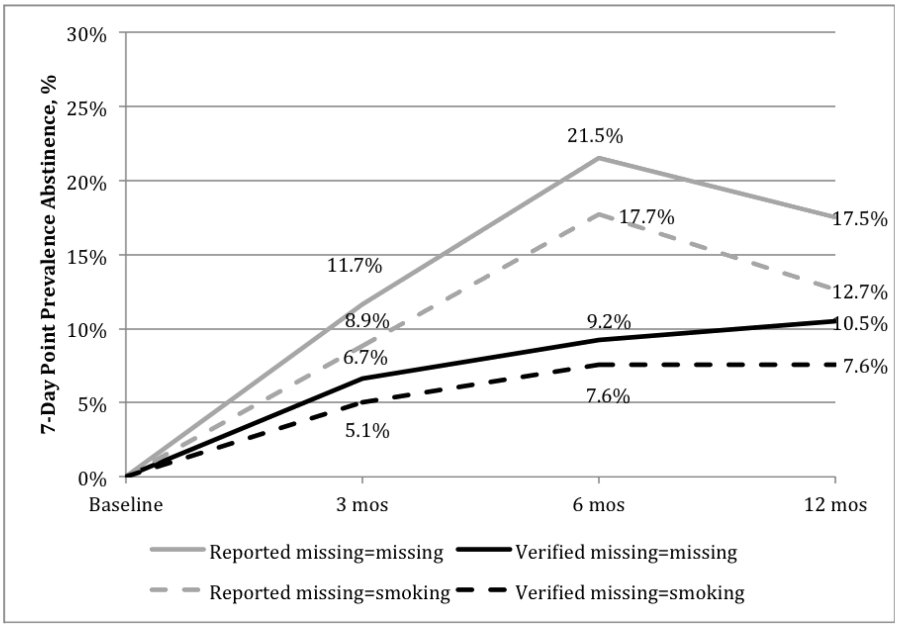


Figure 2. Reported and biochemically verified 7-day point prevalence abstinence by time in a Facebook smoking cessation intervention among young adults in an intent-to-treat (missing=smoking) and complete case (missing=missing) analysis (N=79). Follow-up rates are 76% at 3 months, 82% at 6 months, and 72% at 12 months. Verified abstinence includes only those who returned biochemical verification of abstinence at each follow-up assessment.



Secondary Smoking Outcomes

Quit Attempts, Cessation Aids, and Smoking Reduction

Between baseline and 12-month follow-up, 66% (52/79) of the sample reported at least one purposeful quit attempt lasting at least 24 hours: 46% (36/79) at 3 months, 40% (32/79) at 6 months, 39% (31/79) at 12 months. During the 12 months, in addition to support through the Tobacco Status Project, 11% (9/79) of participants used a nicotine replacement therapy approved by the Food and Drug Administration (5% nicotine gum, 5% nicotine patch, 1% nicotine inhaler), and 18% (14/79) reported using an electronic nicotine delivery system (ENDS) to help them quit smoking. Almost half (38/79, 48%) reported reducing the number of cigarettes smoked by 50% or greater from baseline to 3 and 6 months, and 35% (28/79) reduced at least 50% from baseline to 12 months.

Transitions in Stage of Change

More participants were in preparation or action at intervention-end (3-month follow-up; 32/79, 40%) than at baseline (10/79, 13%; Bowker's X^2=66.7, *P*<.001). At 6 months, more participants were quit or ready to quit (37/79, 47%) than at 3 months (32/79, 40%; Bowker's X^2=66.7, *P*<.001). Transitions from 6-12 months (36/79, 46% ready to quit) were not significant (Bowker's X^2=5.03, *P*=.754).

Engagement

Number of Likes and Comments

Half of participants (40/79, 51%) ‘‘liked’’ at least one study-related post on the Facebook group. The median number of “likes” per person among those 40 participants was 4.0: interquartile range (IQR) 5.5; range 1-73. Almost two-thirds of participants (48/79, 61%) commented on at least one post. Two participants commented on all 90 daily postings. The median number of comments per person among those who posted or commented was 12.0 (IQR 19.5; range 1-78). Volume of commenting did not significantly differ by group, showing Precontemplation: median 0 (IQR 11); Contemplation: median 4.5 (IQR 10.8); Preparation: median 9.5 (IQR 30.5); U_2=5.06, *P*=.080. With the full sample, comment volume was not significantly related to biochemically verified abstinence status at 3 months (U=417.5, *P*=.103). Yet among those who commented at least once (n=48), volume of commenting was significantly greater among those who quit (median 49) compared to those who did not: median 8.0 (IQR 15.0); U=115.50, *P*=.036. Comment count was not significantly related to reduction of cigarettes smoked by 50% or greater, having made a quit attempt during the study period, or readiness to quit smoking at the 12-month follow-up.

Incentive Effects

For the full sample (N=79), there was no significant difference among incentive conditions on number of comments made to Facebook groups: personal median 8 (IQR 19); altruistic median 2 (IQR 6); no incentive median 2.5 (IQR 11.8); $\chi^2_{2,79}$=.035, *P*=.749. Among those who commented at least once (48/79, 61%), those in the personal incentive condition made more comments than those in the other two conditions: personal median 16 (IQR 24); altruistic median 5.5 (IQR 12.3); no incentive median 7 (IQR 12); $\chi^2_{2,79}$=8.44, *P*=.015. There was no significant relationship between altruism and comments within the altruistic incentive group (*P*=.99). There were no difference by incentive condition on reported 3-month abstinence rates ($\chi^2_{2,60}$=3.17, *P*=.205) or biochemically verified 3-month abstinence ($\chi^2_{2,60}$=1.80, *P*=.406).

Usability and Satisfaction With Intervention

More than a third of participants (62/79, 35%) reported reading most or all of the Facebook posts, and 24% (19/79) read most or all of “The Dr. Is In” sessions (Figure 3). Highest ratings were for post ease (57/79, 72%), thinking about what they read (52/79, 66%), and would recommend the program to others (56/79, 71%). Among the 22 CBT counseling sessions rated by 5 participants, all were rated as easy to understand, providing sound advice, and would recommend the program to others; 90% (20/22) of sessions were rated helpful; 90% (20/22) referred to material after the session; and 82% (18/22) of sessions had information that participants later used to make a behavior change (Figure 4).

Figure 3. Proportion of respondents reporting they “agree” or “strongly agree” with statements about their Facebook group for a Facebook smoking cessation intervention (N=62).

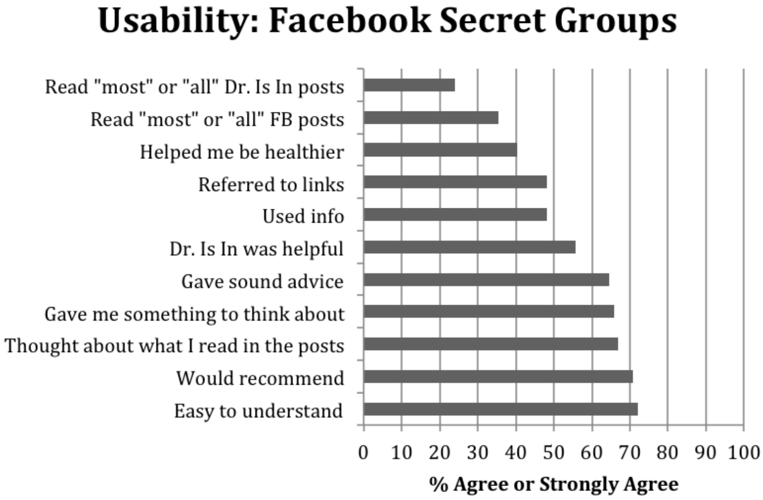
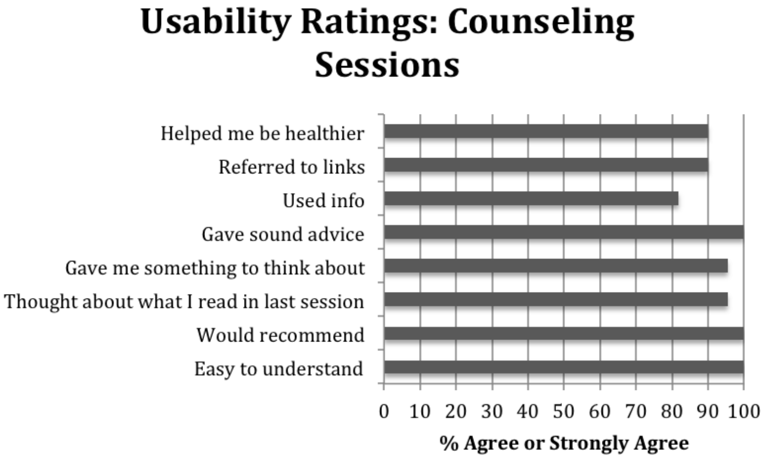


Figure 4. Proportion of “agree” or “strongly agree” reports about 7 cognitive-behavioral counseling sessions for a Facebook smoking cessation intervention (22 session reports).



Discussion

Principal Findings

In our pilot investigation, a novel Facebook-integrated smoking cessation intervention was found to be likable and feasible to deliver to young adults who smoke cigarettes. Early efficacy data are encouraging and support further investigation. Retention also was good, especially for a digital health intervention, though it dropped off for those in preparation at 12 months. Unknown is whether the loss to follow-up was related to quitting or relapsing. This was the smallest group, hence, the values are less stable and more subject to variability. The low sample size in this group could have led to less engagement with the intervention, subsequently leading to a lower desire to participate in follow-up surveys. However, usability data did not differ across readiness to quit groups. A trial with a sample size to support larger groups of those ready to quit is warranted.

Reported and biochemically verified abstinence increased from 3 months to 6 months and then declined slightly at 12 months, while biochemically verified abstinence was flat between 6 and 12 months. Previous trials testing the efficacy of stage-tailored interventions for smoking cessation and other health risk behaviors have found increased abstinence rates over 24 months with repeated exposure to treatment across 1 year (eg, [50]). Participants had access to Facebook groups and all associated materials over the 12-month trial; however, extending the active intervention phase (daily postings) past 3 months, or having “booster” postings over time, may enhance the efficacy of the intervention at 12 months.

Monetary incentives increased intervention uptake, but only for those who had some level of engagement (ie, commented at least once during the 3-month intervention). Further, engagement was related to abstinence only among those who commented at least once. For those motivated to engage, monetary incentives appear to be an important adjunct to a health behavior change intervention and could translate to behavior change. Monetary incentives have been identified as a main motivator for participation in online mental health interventions [51]. For smoking cessation interventions with young people, incentives have been associated with lower retention [52]. In our study, we tied a single monetary incentive (US $50) to commenting on all daily posts for 90 days. Findings suggest that a personal incentive was more effective than one in which participants provided it to others regardless of trait altruism. A more frequent incentive (eg, weekly) may further increase participation. Future studies should test various incentive schemes for engagement, balance incentives across participants and stages of change, and be powered to test the effects of groups to better determine the strategies most effective to increase engagement and for whom.

Participation in CBT counseling was relatively low, yet attendance among those who elected to receive CBT was high (6 of 7 sessions were attended on average). Conducting counseling sessions through Facebook chat was thought to provide a more personal connection. To maximize engagement, however, these sessions could be offered in a more open format than invited chat sessions, perhaps through the stage tailored groups to which all intervention participants were assigned. This would allow for those motivated to engage to do so and allow others to still view session material (eg, handouts, tip sheets, interactions between other participants and study counselors). Additionally, allowing participants to move between groups in response to an increasing stage of change was not used by many and potentially may undermine the social support generated from participating in a single group; thus, this strategy should likely not be used in the future.

Our feasibility study found that two-thirds of all Facebook intervention participants and one third of participants in the Precontemplation stage of change undertook one or more quit attempts, with each participant making at least one attempt reporting more than four on average. These findings are in line with results from daily assessments [53] and speak to the dynamic nature of the smoking cessation process and to the fact that many smokers need multiple quit attempts in order to attain smoking abstinence. However, although we used a standard definition of quit attempts (24 hours of purposeful abstinence), our inclusion of non-daily smokers could have led to overreporting of non-smoking days as quit attempts, although the assessment made clear that the definition included only purposeful attempts to quit smoking. Trying to quit without assistance other than the Tobacco Status Project intervention was much more popular than using a cessation aid such as medication among young adults. This is congruent with previous studies among young adults [54] and adolescents [55] and highlights the need to incorporate evidence-based treatment into media that are widely accessed by young people such as Facebook. Electronic cigarettes were used more frequently as a cessation aid than nicotine replacement therapy, even though clear evidence for their effectiveness is still lacking [56,57] and their use was not recommended in the intervention. Young adults may have strong expectations for the effectiveness of electronic cigarettes as a smoking cessation strategy [58]. Given the wide marketing of these products online to young adults [59], the high use in our sample is not surprising. Until convincing data on the effectiveness of electronic cigarettes for smoking cessation is presented, their use should not be recommended to young adults trying to quit smoking.

We tested a novel method of collecting biochemical verification of tobacco abstinence by mailing participants saliva cotinine test strips and asking them to send back pictures of test results. Incentives for assessment completion were not given until proof was sent. This strategy was successful for about half of those reporting abstinence at 3-, 6-, and 12-month assessments. Of the 13 cases for which we were unable to obtain biochemical verification of abstinence at any time point, 6 were unreachable, 1 had technical difficulties preventing transmission of verification data (photos), 1 had difficulty completing the test at home and results were seen as inconclusive, and 5 were not sent test kits due to errors in survey programming on the study end. This piloting of methods was valuable for informing quality control in our future efforts. Overall success with completing the biochemical procedure at home once test kits were sent was promising, and our strategy shows that validation measures are possible in an environment where information is increasingly easy to send as data (eg, through a smartphone). Notably, we did not have trouble recruiting young people into the study despite knowledge that the trial would ask for biochemical verification of abstinence, as has been a concern of others [60,61].

Limitations

With a goal of determining feasibility and initial efficacy, this study was not adequately powered to fully test many of the relationships examined here. A larger, 2-group randomized trial powered to detect effects on primary outcomes and differences among incentive or motivation groups is warranted. Further, Facebook is a dynamic environment where design issues become out-of-date quickly. Potential changes to user agreements mean that researchers must be vigilant about privacy and confidentiality when working in this medium. Formative work with our target population indicated that young adult smokers were ashamed of smoking in many cases and wanted to keep the study involvement private from their larger Facebook social network [23]. This suggested the intervention should be conducted within the context of Facebook secret groups where only study administrators and group members have knowledge of their existence. Investigators then watched Facebook's privacy practices to ensure that this did not change; if so, we would have considered a different context in which to run groups.

Another limitation is the general lack of sample diversity. In contrast to other online interventions in the United States or Canada that show a majority of female participants [62,63], our sample was primarily male. This was a surprise, given that a larger proportion of American women online use Facebook than men (77% vs 66%) [64]. There was no direct targeting of participants by gender in the Facebook advertising campaign, as one goal of this feasibility study was to determine the characteristics of participants who would enroll. Strategies are needed to recruit more female and ethnic minority participants. Facebook targeting can be used to place ads in locations including states and cities where more ethnic minority smokers reside. Advertising images can be used to target women and non-white smokers. Bull et al [65] recruited a large proportion of non-white participants into a Facebook sexual health intervention using respondent-driven sampling. Using a strategy to recruit "seed" participants and friends in their (real-world) social networks could generate samples that are diverse with respect to gender, ethnicity, and sexual orientation.

Conclusions

Harnessing the popularity of social media to treat tobacco use in young adults holds great potential considering the overwhelming numbers using this medium daily. Focused on young adult smokers, a challenging group to engage, our study's high retention and usability ratings suggest the Facebook quit smoking intervention is attractive and feasible to deliver. Early efficacy data are encouraging and support further investigation in a larger sample with a randomized design.

Acknowledgments

This work was supported by the National Institute on Drug Abuse (NIDA) K23 DA032578 and P50 DA09253. Preparation of this manuscript was supported by NIDA K23 DA018691, the National Cancer Institute R25 CA113710 and the National Institute of Mental Health R01 MH083684, the State of California Tobacco-Related Disease Research Program 21BT-0018, and the National Heart, Lung, and Blood Institute T32 HL007034-37.

Authors' Contributions

DR designed the parent study and wrote the protocol in consultation with JP and KD. DR and KC primarily carried out the study protocol. DR and JT conducted the analyses in consultation with KD and JP. DR, JT, and KC completed the first draft of the manuscript, including all parts, and all authors reviewed and revised subsequent drafts of the manuscript. All authors contributed to and have approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Altruism measure.

[PDF File (Adobe PDF File), 55KB - jmir_v17i12e291_app1.pdf]

References

1. Substance Abuse Mental Health Services Administration. Table 224B - Cigarette Use in Lifetime, Past Year, Past Month among Persons Aged 18 to 25, by Demographic Characteristics: Percentages. In: Results from the 2013 National Survey on Drug Use and Health. Rockville, MD: Department of Health and Human Services; 2013.
2. Curry S, Sporer AK, Pugach O, Campbell RT, Emery S. Use of tobacco cessation treatments among young adult smokers: 2005 National Health Interview Survey. Am J Public Health 2007 Aug;97(8):1464-1469. [doi: 10.2105/AJPH.2006.103788] [Medline: 17600243]
3. Bost ML. A descriptive study of barriers to enrollment in a collegiate Health Assessment Program. J Community Health Nurs 2005;22(1):15-22. [doi: 10.1207/s15327655jchn2201_2] [Medline: 15695193]
4. Davies J, McCrae BP, Frank J, Dochnahl A, Pickering T, Harrison B, et al. Identifying male college students' perceived health needs, barriers to seeking help, and recommendations to help men adopt healthier lifestyles. J Am Coll Health 2000 May;48(6):259-267. [doi: 10.1080/07448480009596267] [Medline: 10863869]
5. Cohen S, Lemay EP. Why would social networks be linked to affect and health practices? Health Psychol 2007 Jul;26(4):410-417. [doi: 10.1037/0278-6133.26.4.410] [Medline: 17605560]
6. Huang GC, Unger JB, Soto D, Fujimoto K, Pentz MA, Jordan-Marsh M, et al. Peer influences: the impact of online and offline friendship networks on adolescent smoking and alcohol use. J Adolesc Health 2014 May;54(5):508-514. [doi: 10.1016/j.jadohealth.2013.07.001] [Medline: 24012065]
7. Burri M, Baujard V, Etter J. A qualitative analysis of an internet discussion forum for recent ex-smokers. Nicotine Tob Res 2006 Dec;8 Suppl 1:S13-S19. [Medline: 17491166]
8. van MT, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. J Med Internet Res 2012;14(3):e66 [FREE Full text] [doi: 10.2196/jmir.1854] [Medline: 22732103]
9. Alexa. Top sites in United States. 2015. URL: http://www.alexa.com/siteinfo/facebook.com [accessed 2015-08-10] [WebCite Cache ID 6dsWvpZSX]
10. statistica. 2015. Number of Facebook users in the United States as of January 2015, by age group (in millions) URL: http://www.statista.com/statistics/398136/us-facebook-user-age-groups/ [accessed 2015-10-07] [WebCite Cache ID 6c6lesrTH]

11. Duggan M. Social Media Update. Washington, DC: Pew Research Center; 2015. URL: http://www.webcitation.org/6dsPQJhqo [accessed 2015-09-16]
12. Bull SS, Levine DK, Black SR, Schmiege SJ, Santelli J. Social media-delivered sexual health intervention: a cluster randomized controlled trial. Am J Prev Med 2012 Nov;43(5):467-474 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.022] [Medline: 23079168]
13. Patrick K, Marshall SJ, Davila EP, Kolodziejczyk JK, Fowler JH, Calfas KJ, et al. Design and implementation of a randomized controlled social and mobile weight loss trial for young adults (project SMART). Contemp Clin Trials 2014 Jan;37(1):10-18 [FREE Full text] [doi: 10.1016/j.cct.2013.11.001] [Medline: 24215774]
14. Young SD. Social media technologies for HIV prevention study retention among minority men who have sex with men (MSM). AIDS Behav 2014 Sep;18(9):1625-1629 [FREE Full text] [doi: 10.1007/s10461-013-0604-z] [Medline: 24062015]
15. Kernot J, Olds T, Lewis LK, Maher C. Effectiveness of a facebook-delivered physical activity intervention for post-partum women: a randomized controlled trial protocol. BMC Public Health 2013;13:518 [FREE Full text] [doi: 10.1186/1471-2458-13-518] [Medline: 23714411]
16. Cavallo DN, Tate DF, Ries AV, Brown JD, DeVellis RF, Ammerman AS. A social media-based physical activity intervention: a randomized controlled trial. Am J Prev Med 2012 Nov;43(5):527-532 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.019] [Medline: 23079176]
17. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]
18. Pedrana A, Hellard M, Gold J, Ata N, Chang S, Howard S, et al. Queer as F**k: reaching and engaging gay men in sexual health promotion through social networking sites. J Med Internet Res 2013;15(2):e25 [FREE Full text] [doi: 10.2196/jmir.2334] [Medline: 23391459]
19. Williams G, Hamm MP, Shulhan J, Vandermeer B, Hartling L. Social media interventions for diet and exercise behaviours: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2014;4(2):e003926 [FREE Full text] [doi: 10.1136/bmjopen-2013-003926] [Medline: 24525388]
20. Jones K, Baldwin KA, Lewis PR. The potential influence of a social media intervention on risky sexual behavior and Chlamydia incidence. J Community Health Nurs 2012;29(2):106-120. [doi: 10.1080/07370016.2012.670579] [Medline: 22536914]
21. Valle CG, Tate DF, Mayer DK, Allicock M, Cai J. Exploring Mediators of Physical Activity in Young Adult Cancer Survivors: Evidence from a Randomized Trial of a Facebook-Based Physical Activity Intervention. J Adolesc Young Adult Oncol 2015 Mar 1;4(1):26-33. [doi: 10.1089/jayao.2014.0034] [Medline: 25852972]
22. Baskerville NB, Azagba S, Norman C, McKeown K, Brown KS. Effect of a Digital Social Media Campaign on Young Adult Smoking Cessation. Nicotine Tob Res 2015 Jun 4:-. [doi: 10.1093/ntr/ntv119] [Medline: 26045252]
23. Ramo DE, Liu H, Prochaska JJ. A mixed-methods study of young adults' receptivity to using Facebook for smoking cessation: if you build it, will they come? Am J Health Promot 2015;29(4):e126-e135. [doi: 10.4278/ajhp.130326-QUAL-128] [Medline: 24575728]
24. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39 [FREE Full text] [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
25. Kohl LFM, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]
26. Caldwell PHY, Hamilton S, Tan A, Craig JC. Strategies for increasing recruitment to randomised controlled trials: systematic review. PLoS Med 2010;7(11):e1000368 [FREE Full text] [doi: 10.1371/journal.pmed.1000368] [Medline: 21085696]
27. Benishek LA, Dugosh KL, Kirby KC, Matejkowski J, Clements NT, Seymour BL, et al. Prize-based contingency management for the treatment of substance abusers: a meta-analysis. Addiction 2014 Sep;109(9):1426-1436 [FREE Full text] [doi: 10.1111/add.12589] [Medline: 24750232]
28. Meier S. A Survey on Ecomonic Theories and Field Evidence on Pro-social Behavior. In: Economics and psychology: a promising new cross-disciplinary field. Cambridge, MA: MIT Press; 2007:51-88.
29. Centers for Disease ControlPrevention (CDC). Vital signs: current cigarette smoking among adults aged >or=18 years --- United States, 2009. MMWR Morb Mortal Wkly Rep 2010 Sep 10;59(35):1135-1140 [FREE Full text] [Medline: 20829747]
30. Ramo DE, Rodriguez TMS, Chavez K, Sommer MJ, Prochaska JJ. Facebook Recruitment of Young Adult Smokers for a Cessation Trial: Methods, Metrics, and Lessons Learned. Internet Interv 2014 Apr;1(2):58-64 [FREE Full text] [doi: 10.1016/j.invent.2014.05.001] [Medline: 25045624]
31. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 6863699]
32. DiClemente CC, Prochaska JO, Fairhurst SK, Velicer WF, Velasquez MM, Rossi JS. The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change. J Consult Clin Psychol 1991 Apr;59(2):295-304. [Medline: 2030191]
33. Center F. What are the privacy options for groups?. 2015. H URL: https://www.facebook.com/unsupportedbrowser [accessed 2015-12-18] [WebCite Cache ID 6dsQJVgav]

34. Clinical Practice Guideline Treating Tobacco Use and Dependence 2008 Update Panel, Liaisons,Staff. A clinical practice guideline for treating tobacco use and dependence: 2008 update. A U.S. Public Health Service report. Am J Prev Med 2008 Aug;35(2):158-176 [FREE Full text] [doi: 10.1016/j.amepre.2008.04.009] [Medline: 18617085]
35. Carpenter MJ, Hughes JR, Solomon LJ, Callas PW. Both smoking reduction with nicotine replacement therapy and motivational advice increase future cessation among smokers unmotivated to quit. J Consult Clin Psychol 2004 Jun;72(3):371-381. [doi: 10.1037/0022-006X.72.3.371] [Medline: 15279521]
36. Carpenter MJ, Hughes JR, Keely JP. Effect of smoking reduction on later cessation: a pilot experimental study. Nicotine Tob Res 2003 Apr;5(2):155-162. [Medline: 12745487]
37. Chan SC, Lam TH, Salili F, Leung GM, Wong DCN, Botelho RJ, et al. A randomized controlled trial of an individualized motivational intervention on smoking cessation for parents of sick children: a pilot study. Appl Nurs Res 2005 Aug;18(3):178-181. [doi: 10.1016/j.apnr.2005.01.002] [Medline: 16106336]
38. Butler C, Rollnick S, Cohen D, Bachmann M, Russell I, Stott N. Motivational consulting versus brief advice for smokers in general practice. Br J Gen Pract 1999;49:611-616.
39. Curry SJ, Ludman EJ, Graham E, Stout J, Grothaus L, Lozano P. Pediatric-based smoking cessation intervention for low-income women: a randomized trial. Arch Pediatr Adolesc Med 2003 Mar;157(3):295-302. [Medline: 12622686]
40. Brown R, Ramsey SE, Strong DR, Myers MG, Kahler CW, Lejuez CW, et al. Effects of motivational interviewing on smoking cessation in adolescents with psychiatric disorders. Tob Control 2003 Dec;12 Suppl 4:IV3-I10 [FREE Full text] [Medline: 14645934]
41. Steinberg ML, Ziedonis DM, Krejci JA, Brandon TH. Motivational interviewing with personalized feedback: a brief intervention for motivating smokers with schizophrenia to seek treatment for tobacco dependence. J Consult Clin Psychol 2004 Aug;72(4):723-728. [doi: 10.1037/0022-006X.72.4.723] [Medline: 15301657]
42. Pro-Change Behavior Systems. A Guide for Smoking Cessation. South Kingstown, RI: Pro-Change Behavior Systems; 2009. URL: http://www.prochange.com/smoking-cessation-program [accessed 2015-12-22] [WebCite Cache ID 6dyI8AgcR]
43. Hettema JE, Hendricks PS. Motivational interviewing for smoking cessation: a meta-analytic review. J Consult Clin Psychol 2010 Dec;78(6):868-884. [doi: 10.1037/a0021498] [Medline: 21114344]
44. Hall SM, Tsoh JY, Prochaska JJ, Eisendrath S, Rossi JS, Redding CA, et al. Treatment for cigarette smoking among depressed mental health outpatients: a randomized clinical trial. Am J Public Health 2006 Oct;96(10):1808-1814. [doi: 10.2105/AJPH.2005.080382] [Medline: 17008577]
45. Thrul J, Stemmler M, Bühler A, Goecke M. The role of participants' self-selected future smoking goals in adolescent smoking cessation interventions. Drug Alcohol Depend 2014 Aug 1;141:118-123. [doi: 10.1016/j.drugalcdep.2014.05.016] [Medline: 24928478]
46. Transdisciplinary Tobacco Use Research Center (TTURC) Tobacco Dependence, Baker TB, Piper ME, McCarthy DE, Bolt DM, Smith SS, et al. Time to first cigarette in the morning as an index of ability to quit smoking: implications for nicotine dependence. Nicotine Tob Res 2007 Nov;9 Suppl 4:S555-S570 [FREE Full text] [doi: 10.1080/14622200701673480] [Medline: 18067032]
47. Hughes JR, Keely JP, Niaura RS, Ossip-Klein DJ, Richmond RL, Swan GE. Measures of abstinence in clinical trials: issues and recommendations. Nicotine Tob Res 2003 Feb;5(1):13-25. [Medline: 12745503]
48. Andreoni J, Miller J. Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism. Econometrica 2002 Mar;70(2):737-753. [doi: 10.1111/1468-0262.00302]
49. Ben-Ner A, Kramer A, Levy O. Economic and hypothetical dictator game experiments: Incentive effects at the individual level. J Socio-Econ 2008 Oct;37(5):1775-1784. [doi: 10.1016/j.socec.2007.11.004]
50. Prochaska JO, Velicer WF, Redding C, Rossi JS, Goldstein M, DePue J, et al. Stage-based expert systems to guide a population of primary care patients to quit smoking, eat healthier, prevent skin cancer, and receive regular mammograms. Prev Med 2005 Aug;41(2):406-416. [doi: 10.1016/j.ypmed.2004.09.050] [Medline: 15896835]
51. Crisp DA, Griffiths KM. Participating in online mental health interventions: who is most likely to sign up and why? Depress Res Treat 2014;2014:790457 [FREE Full text] [doi: 10.1155/2014/790457] [Medline: 24804089]
52. Backinger CL, Michaels CN, Jefferson AM, Fagan P, Hurd AL, Grana R. Factors associated with recruitment and retention of youth into smoking cessation intervention studies--a review of the literature. Health Educ Res 2008 Apr;23(2):359-368 [FREE Full text] [doi: 10.1093/her/cym053] [Medline: 17884834]
53. Hughes JR, Solomon LJ, Naud S, Fingar JR, Helzer JE, Callas PW. Natural history of attempts to stop smoking. Nicotine Tob Res 2014 Sep;16(9):1190-1198 [FREE Full text] [doi: 10.1093/ntr/ntu052] [Medline: 24719491]
54. Solberg LI, Asche SE, Boyle R, McCarty MC, Thoele MJ. Smoking and cessation behaviors among young adults of various educational backgrounds. Am J Public Health 2007 Aug;97(8):1421-1426. [doi: 10.2105/AJPH.2006.098491] [Medline: 17600256]
55. Myers MG, MacPherson L. Smoking cessation efforts among substance abusing adolescents. Drug Alcohol Depend 2004 Feb 7;73(2):209-213. [Medline: 14725961]
56. Grana R, Benowitz N, Glantz SA. E-cigarettes: a scientific review. Circulation 2014 May 13;129(19):1972-1986 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.114.007667] [Medline: 24821826]

57. Hajek P, Foulds J, Le HJ, Sweanor D, Yach D. Should e-cigarettes be regulated as a medicinal device? Lancet Respir Med 2013 Aug;1(6):429-431. [doi: 10.1016/S2213-2600(13)70124-3] [Medline: 24429230]
58. Choi K, Fabian L, Mottey N, Corbett A, Forster J. Young adults' favorable perceptions of snus, dissolvable tobacco products, and electronic cigarettes: findings from a focus group study. Am J Public Health 2012 Nov;102(11):2088-2093 [FREE Full text] [doi: 10.2105/AJPH.2011.300525] [Medline: 22813086]
59. Grana RA, Ling PM. "Smoking revolution": a content analysis of electronic cigarette retail websites. Am J Prev Med 2014 Apr;46(4):395-403 [FREE Full text] [doi: 10.1016/j.amepre.2013.12.010] [Medline: 24650842]
60. Stanton A, Grimshaw G. Tobacco cessation interventions for young people. Cochrane Database Syst Rev 2013;8:CD003289. [doi: 10.1002/14651858.CD003289.pub5] [Medline: 23975659]
61. Danaher BG, Severson HH, Andrews JA, Tyler MS, Lichtenstein E, Woolley TG, et al. Randomized controlled trial of MyLastDip: a Web-based smokeless tobacco cessation program for chewers ages 14-25. Nicotine Tob Res 2013 Sep;15(9):1502-1510 [FREE Full text] [doi: 10.1093/ntr/ntt006] [Medline: 23410803]
62. Pechmann C, Pan L, Delucchi K, Lakon CM, Prochaska JJ. Development of a Twitter-based intervention for smoking cessation that encourages high-quality social media interactions via automessages. J Med Internet Res 2015;17(2):e50 [FREE Full text] [doi: 10.2196/jmir.3772] [Medline: 25707037]
63. Nash CM, Vickerman KA, Kellogg ES, Zbikowski SM. Utilization of a Web-based vs integrated phone/Web cessation program among 140,000 tobacco users: an evaluation across 10 free state quitlines. J Med Internet Res 2015;17(2):e36 [FREE Full text] [doi: 10.2196/jmir.3658] [Medline: 25673013]
64. Pew Internet & American Life Project. Pew Research Center's Internet & American Life Project. 2015. Facebook Demographics URL: http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/2015-08-19_social-media-update_07/ [accessed 2015-11-03] [WebCite Cache ID 6dsRxei2t]
65. Bull SS, Breslin LT, Wright EE, Black SR, Levine D, Santelli JS. Case study: An ethics case study of HIV prevention research on Facebook: the Just/Us study. J Pediatr Psychol 2011;36(10):1082-1092 [FREE Full text] [doi: 10.1093/jpepsy/jsq126] [Medline: 21292724]

Abbreviations

ANOVA: analysis of variance
CBT: cognitive behavioral therapy
ENDS: electronic nicotine delivery systems
FDA: Food and Drug Administration
HIPAA: Health Insurance Portability and Accountability Act
IQR: interquartile range

Edited by G Eysenbach; submitted 07.10.15; peer-reviewed by G Norman, RS Sadasivam, M Mehring; comments to author 29.10.15; revised version received 09.11.15; accepted 14.11.15; published 31.12.15.

Please cite as:
Ramo DE, Thrul J, Chavez K, Delucchi KL, Prochaska JJ
Feasibility and Quit Rates of the Tobacco Status Project: A Facebook Smoking Cessation Intervention for Young Adults
J Med Internet Res 2015;17(12):e291
URL: http://www.jmir.org/2015/12/e291/
doi:10.2196/jmir.5209
PMID:26721211

Original Paper

Apps and Adolescents: A Systematic Review of Adolescents' Use of Mobile Phone and Tablet Apps That Support Personal Management of Their Chronic or Long-Term Physical Conditions

Rabiya Majeed-Ariss[1], BSc (Hons), MSc, PhD; Eileen Baildam[2], MB, ChB, FRCPCH; Malcolm Campbell[3], BSc (Hons), MSc, PhD; Alice Chieng[4], MBCHB, MSc, MRCPaed; Debbie Fallon[3], BSc, MA, PhD; Andrew Hall[3], BA (Hons), MEd, FHEA; Janet E McDonagh[5,6], MBBS, MD, FRCP; Simon R Stones[7]; Wendy Thomson[5,6], BSc, PhD; Veronica Swallow[8], BSc (Hons), MMedSc, PhD

[1]University of Manchester, School of Psychological Sciences, Manchester, United Kingdom

[2]Alder Hey Children's Foundation National Health Service (NHS) Foundation Trust, Department of Rheumatology, Liverpool, United Kingdom

[3]University of Manchester, School of Nursing, Midwifery, and Social Work, Manchester, United Kingdom

[4]Royal Manchester Children's Hospital, Department of Rheumatology, Manchester, United Kingdom

[5]Arthritis Research UK Centre for Epidemiology, Centre for Musculoskeletal Research, University of Manchester, Manchester, United Kingdom

[6]National Institute for Health Research (NIHR) Manchester Musculoskeletal Biomedical Research Unit, Manchester Academic Health Science Centre, University of Manchester, Manchester, United Kingdom

[7]The University of Manchester, Faculty of Life Sciences, Manchester, United Kingdom

[8]University of Leeds, Leeds, United Kingdom

Corresponding Author:
Veronica Swallow, BSc (Hons), MMedSc, PhD
University of Leeds
School of Healthcare, Baines Wing, University of Leeds
Leeds, LS2 9JT
United Kingdom
Phone: 44 0113 243 1751
Fax: 44 0113 243 1753
Email: v.m.swallow@leeds.ac.uk

Abstract

Background: The prevalence of physical chronic or long-term conditions in adolescents aged 10-24 years is rising. Mobile phone and tablet mobile technologies featuring software program apps are widely used by these adolescents and their healthy peers for social networking or gaming. Apps are also used in health care to support personal condition management and they have considerable potential in this context. There is a growing body of literature on app use in health contexts, thereby making a systematic review of their effectiveness very timely.

Objective: To systematically review the literature on the effectiveness of mobile apps designed to support adolescents' management of their physical chronic or long-term conditions.

Methods: We conducted a review of the English-language literature published since 2003 in five relevant bibliographical databases using key search terms. Two independent reviewers screened titles and abstracts using data extraction and quality assessment tools.

Results: The search returned 1120 hits. Of the 19 eligible full-text papers, four met our review criteria, reporting one pilot randomized controlled trial and three pretest/post-test studies. Samples ranged from 4 to 18 participants, with a combined sample of 46 participants. The apps reported were targeted at type 1 diabetes, asthma, and cancer. Two papers provided data for calculating effect size. Heterogeneity in terms of study design, reported outcomes, follow-up times, participants' ages, and health conditions prevented meta-analyses. There was variation in whether adolescents received guidance in using the app or were solely responsible for navigating the app. Three studies reported some level of patient involvement in app design, development, and/or evaluation. Health professional involvement in the modelling stages of apps was reported in all studies, although it was not always clear whether specific clinical (as opposed to academic) expertise in working with adolescents was represented. The dearth of studies and the small overall sample size emphasizes the need for future studies of the development, evaluation, use, and effectiveness of mobile apps to support adolescents' personal management of their conditions.

Conclusions: A key finding of the review is the paucity of evidence-based apps that exist, in contrast to the thousands of apps available on the app market that are not evidence-based or user or professional informed. Although we aimed to assess the effectiveness of apps, the dearth of studies meeting our criteria meant that we were unable to be conclusive in this regard. Based on the available evidence, apps may be considered feasible health interventions, but more studies involving larger sample sizes, and with patient and health professional input at all stages, are needed to determine apps' acceptability and effectiveness. This review provides valuable findings and paves the way for future rigorous development and evaluation of health apps for adolescents with chronic or long-term conditions.

(J Med Internet Res 2015;17(12):e287) doi:10.2196/jmir.5043

KEYWORDS

adolescents; asthma; mobile or tablet apps; arthritis; cancer; chronic disease or condition; diabetes; long-term condition; personal or self-management; review; young people

Introduction

Adolescents With Physical Long-Term or Chronic Conditions

Globally, the pattern of illness in young people or adolescents aged 10-24 years (hereafter referred to as adolescents) has shifted from acute to long-term or chronic conditions (hereafter referred to as chronic). A chronic condition in this age group is a health condition that at the time of diagnosis is predicted to last longer than 3 months [1]. At least 15% of adolescents aged 11-15 years report having been diagnosed with a chronic medical illness or disability [2]. Survival rates for this group have improved due to better screening, earlier detection, and improvements in the delivery of specialized care [1,3,4]. However, there is growing evidence to suggest that young people with chronic conditions have distinct health needs when compared to other groups [4,5].

Effective support from the health sector is therefore paramount, especially during the transition from pediatric to adult health services, and particularly if adult services are not young-person friendly [6]. This process of health transition as young people grow up requires them to develop clinical skills and knowledge in order to ultimately take responsibility for, and competently manage, their personal health care where appropriate [4,7-9]. Delivering safe and timely health care that is accessible and tailored to individuals' needs and preferences is a central feature of international health care strategies [4]. Additionally, government policies highlight the need for services to support self-care; for example, the UK Department of Health and Department for Education are working to support young people with complex health needs in making the transition to adulthood [10].

Contemporaneous reports indicate that utilizing modern mobile electronic technologies in health interventions for young people [11-13] may be a suitable way to address self, shared, or joint care in a manner that is resource efficient.

Significant declines in treatment adherence have been observed during adolescence and the transition from pediatric to adult-centered health services [14]. Education interventions alone are insufficient to promote adherence, but outcomes could be enhanced by adding the following behavioral interventions: monitoring and goal setting, reinforcing medication taking with rewards, contingency contracting, problem solving, and linking medication taking with established routines [15,16]. However, the reported treatment effects are small and reflect the methodological limitations of the included studies and the need to re-examine the delivery and mechanisms of adherence-promoting interventions.

In a recent commentary, Wu and Hommel [17] describe current and potential technologies, such as short message service (SMS) text messaging, mobile phone apps, electronic monitors of adherence, and illness-specific medical devices, to promote pediatric adherence to prescribed medical regimens. The uses reported include the following: delivering and collecting information, facilitating communication between patients and professionals, social networking, capturing real-time data, monitoring bodily functions, automated feedback, guidance and clinical alerts, and smart decision-making tools. However, despite the significant potential and increased use of mobile technologies, to our knowledge there has not been a synthesis of studies reporting on their effectiveness in the management of physical chronic health conditions in adolescents.

Mobile Phone and Tablet Apps to Support Chronic Condition Management

Personal management of chronic physical conditions involves five core skills: problem solving, decision making, resource utilization, forming patient-health professional relationships, and taking action [18]. Apps can support these skills through knowledge development and by providing and collecting information in an accessible, convenient, and interactive way. Mobile phones and tablets form the new generation of mobile electronic devices, different to previous generations in that they are a consumer product as opposed to primarily a business product [19]. Mobile phones and tablets can function with custom software programs called apps, which technologically allow the development of condition-specific and patient-tailored software. These are personal devices, adapted by the user to reflect their specific needs, thus allowing for adaptive, responsive, confidential, and targeted channels of communication and alerts.

A review of the effectiveness of mobile health technology-based health behavior change or disease management interventions for adults found that only six of the 49 interventions used apps and none of these involved adolescents [12]. Another review of mobile phone interventions for management of chronic disease in 18-73-year-olds [20] found few mobile apps and recommended that more be developed. In the Italian health care Android market the majority of apps were designed for health care professionals [21]. Since the potential of mobile technologies in personal health care is significant, a growing body of literature on the use of apps to support patients' management of chronic conditions is emerging.

Mobile apps are widely accepted by adolescents living in today's technology-rich environment. In the United Kingdom for instance, children and adolescents aged 5-15 years are frequent users of mobile technologies. Indeed, 62% of 12-15-year-olds own a mobile phone, and the use of tablet computers by 5-15-year-olds tripled between 2012 and 2013 with 42% using tablets in 2013 [19]; these upward trends are expected to continue. Mobile technologies offer new opportunities to engage adolescents in personal health care [4] but are not without their challenges.

In 2013, the UK National Health Service (NHS) Commissioning Board unveiled a library of NHS-reviewed health apps [22]. Although the review focused on clinical safety rather than clinical effectiveness, it acknowledged that the computing capability contained within mobile technologies offers a legitimate platform for medical and public health practice. However, the Institute of Medical Science (IMS) Institute of Healthcare Informatics (IMS Health) [23] reported that the lack of evidence regarding the effectiveness of mobile apps acts as a barrier to physicians prescribing them. The IMS identified a pressing need for credible evidence of the value of health apps, which in many cases are being used without a thorough understanding of their associated risks and benefits, or a rigorous, evidence-based approach to their development and evaluation [24].

Yet despite increased use and the significant potential of these technologies for adolescents with chronic conditions, to our knowledge a synthesis of studies of their effectiveness in this population has not been undertaken. This systematic review of the evidence is, therefore, timely as it aimed to assess the effectiveness of mobile phone and tablet apps for adolescents' personal management of chronic conditions. In this review, young people are defined as those aged 10-24 years (as defined by the World Health Organization [WHO] [25]) who are undergoing key elements of development, particularly brain development, which continues until the early 20s [5,25,26]. This is arguably a crucial time for the introduction of interventions that promote shared and self-management skills and knowledge, and for the development of both health-promoting as well as health-risk behaviors. The review protocol was published previously in JMIR Research Protocols [27], but key details are reiterated here for new readers.

Methods

The Systematic Review

This systematic review aimed to synthesize the evidence on mobile phone and tablet apps. The methodology adhered to that described in the Cochrane Handbook for Systematic Reviews of Intervention [28] and complies with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [29]. This review was registered with the international prospective register of systematic reviews (PROSPERO) (CRD42014015418) [30].

Search Strategy

Eligible studies were identified through a comprehensive literature search of the following five bibliographical databases: MEDLINE, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), Embase, PsycINFO, and the Web of Science. The search strategy, which was developed in consultation with an information scientist, used standardized indexed search terms and free-text terms that relate to the following three key concepts: (1) adolescents, (2) physical chronic conditions, and (3) mobile technology. The search included British and North American terms and spellings. The search strategy was initially devised in MEDLINE and then adapted to the other databases. The Web of Science did not employ any indexed search terms and the other databases did not employ them in a standardized fashion. Free-text terms were used consistently throughout. In addition to testing search sensitivity, journals associated with the most retrieved citations were hand searched from 2009 to 2014, thus decreasing the likelihood of missing relevant studies. The identification of any studies additional to those we had identified from hand searching allowed us to comment on the rigor of the search strategy and the quality of indexing in the bibliographic databases mentioned above. This is a particularly useful strategy in this relatively new domain of mobile technology. Also, due to the emerging nature of mobile technology, the search included conference abstracts published in peer-reviewed journals, and authors were contacted requesting additional related published or unpublished work.

Screening and Selection Criteria

Overview

Initially, all papers were independently scrutinized by two reviewers (MR, AH) using a screening tool that outlined the study inclusion criteria (see Textbox 1). The 782 articles that met this criteria were then divided between two teams of two reviewers—Team 1: MR, VS and Team 2: AH, DF—who further scrutinized the abstracts using the same screening tool. Whenever disagreement in interpretation arose within one team, the other team was asked to consult the relevant materials to enable a discussion until a consensus between both teams was reached, thereby minimizing bias in the interpretation of findings. Team meetings were held regularly to discuss any complications or challenges.

Inclusion Criteria

Criteria for included studies are shown in Textbox 1.

Textbox 1. Summary of inclusion criteria.

Inclusion criteria

- Population: Adolescents aged 10-24 years (WHO definition from 2001 [25]) diagnosed with chronic physical conditions in any setting.
- Intervention: Any app for a mobile phone or tablet that could be considered a management intervention (or a component of an intervention) in terms of content and/or delivery. This judgment was based on the five core management skills for chronic physical health conditions, as outlined by Lorig [18].
- Comparisons: Intervention versus usual care *or* intervention variant versus intervention variant *or* pre and post.
- Outcomes: Any physiological, attitudinal, behavioral, or knowledge outcomes.
- Study design: Randomized controlled trial (RCT) *or* controlled clinical trial *or* cohort analytic study *or* case-control study *or* cohort study *or* interrupted time series.

The Cochrane Collaboration excludes nonrandomized controlled trials due to their greater bias, but because this is a relatively new area, we included studies of various designs to systematically collect an overview of the current evidence.

Exclusion Criteria

While international literature was included, non-English-language publications and studies that focused on adolescents with mental health problems, learning disabilities, and/or cognitive impairment were excluded due to resource limitations. Interventions using mobile phone technology only in the context of delivering/receiving text messages or phone calls were also excluded. Given the review focus, the technology context was considered key so we applied a publication start date of 2003. This is the year when 3G networks, which provided the bandwidth required to support advanced mobile Internet apps, were launched in the United Kingdom [31]. By January 2007, 147 wideband code-division multiple access (WCDMA)—the standard found in 3G mobile telecommunications—network operators were delivering commercial services to over 100 million subscribers in 67 countries on all continents [32].

Data Extraction

For every included study, two reviewers extracted relevant data independently. A tool based on the data extraction template for Cochrane reviews [33] was developed to facilitate consistent data extraction and prevent important information from being overlooked. Any discrepancies between reviewers were resolved by discussion with the wider research team. Where required, authors were contacted for clarification or additional information. Completed electronic extraction sheets were kept as part of the audit trail, should they be required at a later stage to enable data checking.

Quality Assessment

The evidence and quality of the papers included in the systematic review were assessed using the Effective Public Health Practice Project (EPHPP) quality assessment tool [34,35]. This method requires a review team with at least one member having methodological expertise, and two members with subject expertise; the team for this review met these criteria. The tool involves six component rating domains: selection bias, study design, confounders, blinding, data collection methods, and withdrawals and dropouts. As with the data extraction stage, each study was scored independently by two reviewers, and any disagreements were resolved through discussion with the wider team.

Data Synthesis and Inter-rater Reliability

Except where indicated, extracted data from the papers were analyzed using IBM SPSS Statistics version 22.0 (IBM Corp). For each paper, inter-rater reliability was estimated for each of the domains, the total score, and the final grade of the EPHPP. Following Armijo-Olivo et al [35], agreement for each domain and the final grade before consensus was estimated using Cohen’s unweighted kappa statistic (κ) [36]. Values were interpreted using the criteria of Altman [36]: κ>.80 was interpreted as very good, .61-.80 good, .41-.60 moderate, .21-.40 fair, and ≤.20 poor agreement. Because the scores for each domain and the final grade were ordinal (1=strong, 2=moderate, 3=weak), Cohen’s weighted kappa was also estimated [36]. Unweighted and weighted kappas and their 95% confidence intervals were estimated using an online calculator on the VassarStats website [37]. Again following Armijo-Olivo et al [35], inter-rater agreement on the overall score across the domains was estimated using an intraclass correlation coefficient (ICC), using a two-way, mixed-model analysis of variance with assessor as the fixed factor and paper as the random factor. Values of the ICC were interpreted against the criteria recommended by Armijo-Olivo et al: ICC≥.75 was interpreted as excellent, .60-.74 good, .40-.59 fair-to-moderate, and ≤.40 poor agreement. Fleiss and Cohen [38] demonstrated the equivalence of weighted kappa and the ICC, so these criteria were also used to interpret weighted kappa.

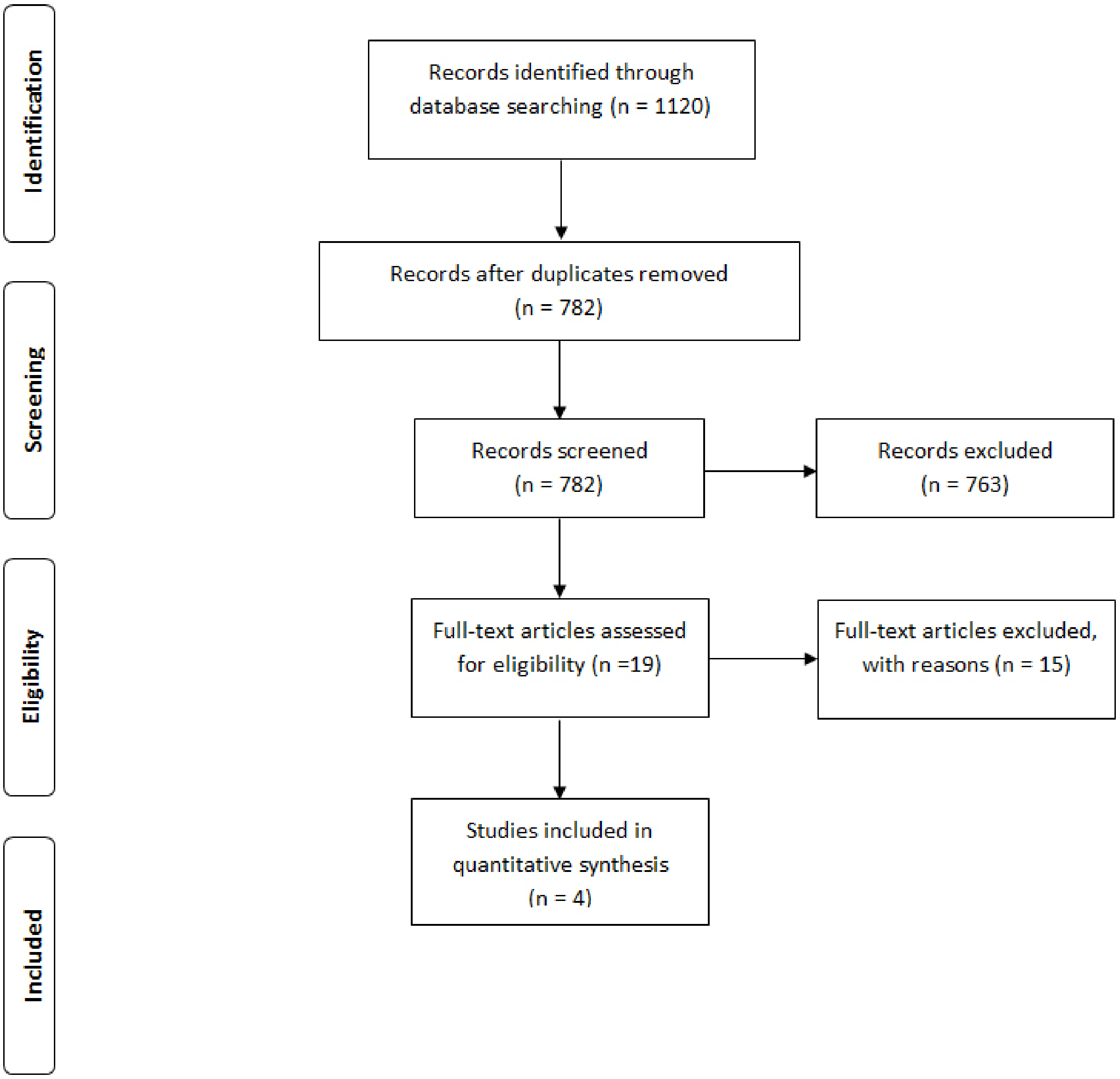
Results

Study Selection

The combined electronic searches identified 1120 records. Of these, 338 records were removed after accounting for duplicates, leaving 782 records for further consideration. Out of the 782 titles and abstracts that were then screened to test eligibility using the PICOS screening tool (population or participant, intervention or indicator, comparator or control, outcome, and study design), 19 full-text papers were potentially eligible for inclusion. Many of the excluded papers reported observational, noncontrolled studies that did not focus on the population or intervention of interest.

Seven of the 19 studies included from the title and abstract screen were abstracts from conference presentations; a search was undertaken to find the full paper of each abstract, failing which the corresponding author was contacted. For two of the seven conference papers, we found subsequent publications [39,40]. We contacted the remaining five authors; three responded with information that meant we were able to exclude their work from the review, and two did not respond meaning their abstracts were also excluded as the full paper was unavailable for consideration in the review. Of the 19 full-text papers considered for eligibility, four papers were assessed as suitable for the full review. Multimedia Appendix 1 lists the respective reasons for excluding the remaining 15 papers [40-54]. Figure 1 illustrates the PRISMA flowchart representing the study selection process. Further hand searching of the Journal of Medical Internet Research from 2009 to 2014 did not identify any additional studies.

Figure 1. PRISMA flow diagram of the review.



Characteristics of Included Studies

Overview

The four included studies described four different apps addressing the needs of adolescents with different chronic conditions: type 1 diabetes [55,56], asthma [39], and cancer [57]. The commonality among all apps studied was that they aimed to support the adolescent in the medical management of their physical condition. Table 1 provides an overview of the four studies. Three of the studies were pretest/post-test designs with no control group [39,55,56] and one was a pilot RCT [57], which used a variety of measurements [58-60]. Two of the studies were conducted in North America [39,55] and two in Western Europe [56,57]. All of the studies recruited adolescents from secondary health care, and follow-up times ranged from 2 to 12 weeks. Based on the data provided, it is not possible to comment on mean age, gender, or ethnicity of the overall sample. The sample sizes ranged from 4 to 18 participants, with a combined total sample of 46 participants. Since the included papers were reporting on feasibility studies with small sample sizes, generalizability of the findings cannot be commented on.

Aldiss et al evaluated the Advanced Symptom Management System (ASyMS) which utilizes mobile phone technology to monitor chemotherapy-related symptoms and promote self-care [57]. This system was first developed with an adult cancer population but Aldiss et al are using a three-phased approach to adapt it for use by adolescents (or young people) with cancer (ASyMS-YG). Phase 1 had involved adolescents identifying the symptoms to be assessed; in Phase 2 adolescents tested the symptom report system, and adolescents', parents', and professionals' perceptions of ASyMS-YG were ascertained [61,62]. The paper included in this review reports on Phase 3 in which the system was developed further in preparation for an RCT. Aldiss et al evaluated it with a group of adolescents with cancer, asking them to complete the ASyMS questionnaire once a day for 14 days during a cycle of chemotherapy.

Table 1. Characteristics of the included studies.

Study	Country	SS[a], n	Measurements	ID[b], weeks	F[c], %	Age (years), mean (range)	Study design	Cond.[d]
Aldiss et al [57]	UK	4	Pediatric Quality of Life Inventory Cancer Module Teen Report Form [58] The Life Situation Scale for Adolescents (LSS-A) [59] The State-Trait Anxiety Inventory (STAI) [60] Perceptions of Technology Questionnaire	2	0	N/A[e] (13-15)	Pilot RCT[f]	Cancer
Burbank et al [39]	USA	18	Self-efficacy questionnaire Asthma Control Test	8	N/A	13.5 (12-17)	Feasibility pretest/post-test	Asthma
Cafazzo et al [55]	Canada	12	Self-care inventory Diabetes Family Responsibility Questionnaire Diabetes QOL[g]Instrument for Youth	12	67	14.9 (12-16)	Pretest/post-test (mixed method)	Type 1 diabetes
Frøisland et al [56]	Norway	12	HbA1c[h] Knowledge test score, before and after the intervention System Usability Scale after the intervention	12	54	16.2 (13-19)	Pretest/post-test (mixed method)	Type 1 diabetes

[a] SS: sample size.

[b] ID: intervention duration.

[c] F: female.

[d] Cond.: condition.

[e] N/A: not applicable (information not reported or stated).

[f] RCT: randomized controlled trial.

[g] QOL: quality of life.

[h] HbA1c: hemoglobin A1c (glycated hemoglobin).

In light of the national asthma education program recommendation for a written asthma action plan (AAP) for all patients with asthma, and in recognition that few studies demonstrating acceptance of phone-based self-monitoring have taken place in rural and medically underserved US regions, Burbank et al tested a mobile AAP mobile phone app for adolescents with persistent asthma in Arkansas [39]. The app was designed to help self-monitoring by recording behaviors as well as prompting positive behaviors by providing immediate feedback on data entered.

Cafazzo et al [55] designed an mHealth intervention for the management of type 1 diabetes in adolescents that aimed to increase the frequency of daily blood glucose readings and to assist with self-care behaviors, establishing the use of technology to improve glycemic control among adolescents as a long-term objective [55]. Addressing the adolescent's need for fast, discrete transfer of blood glucose data, this team developed the bluglu adapter to facilitate automated data transfers (via Bluetooth) from a glucometer to an iPhone or iPad touch device running the bant app; the app then analyzes the data to facilitate feedback to the adolescents in real time. Actions were rewarded with iTunes and apps, which introduced the notion of gamification to this intervention. During a 12-week evaluation, 20 diabetic adolescents aged 12-16 years were supplied with the bant app, glucometer, and bluglu. The outcome measure was the average daily frequency of blood glucose measurement during the pilot compared with the preceding 12 weeks. Finally, Frøisland et al [56] tested adolescents' experiences with a diabetes diary known as Diamob, which recorded data before a mandatory consultation with a health professional to discuss the app and reflect on its recordings.

Effect Sizes

Where possible, Cohen's *d* statistic [63] was calculated as an intention-to-treat effect size for outcome measures in each paper following Donker et al [64]. The one study that was a randomized controlled trial [57] did not report any quantitative results as only 3 of the 4 participants completed the trial. The only papers where estimation of effect sizes was possible were pretest/post-test designs, so Cohen's *d* was calculated for each measure as mean post-test score minus mean pretest score divided by the pooled standard deviation [65]. While this approach does not take into account the repeated nature of the data, the alternative approach to divide the mean difference by the standard deviation of the difference score [65] requires statistical information that is not routinely published. The analysis in one paper [39] was nonparametric, but the authors did not explain why; descriptive statistics for outcomes were reported as medians and interquartile ranges, which were insufficient for reliable estimation of Cohen's *d*. Cohen's *d* was estimated for the remaining two papers using a Microsoft Excel spreadsheet from Missouri State University [66]. Cohen's *d* was interpreted using Cohen's own criteria [63]: 0.80 was interpreted as large, 0.50 as medium, and 0.20 as small.

Cafazzo et al recruited adolescents aged 12-16 years who were diagnosed with type 1 diabetes for more than 1 year, were receiving care at one clinic for at least 6 months, had a hemoglobin A1c (HbA1c) level between 8 and 10%, and were able to communicate in English [55]. Frøisland et al recruited 13-19-year-olds who were diagnosed with type 1 diabetes for more than 1 year, were receiving care at one of two pediatric clinics, and had an HbA1c level of less than 10% [56]. Aldiss et al [57] and Burbank et al [39] did not provide information regarding participant selection and inclusion criteria other than their selection of adolescents with the named chronic condition.

The Interventions

All four apps focused on management of a chronic condition. Cafazzo et al's [55] and Frøisland et al's [56] apps for type 1 diabetes management focused on increasing the number of blood glucose readings. Burbank et al developed an app outlining an AAP to improve asthma management [39]. These three apps can be seen as aids to prevent deterioration of the respective conditions, and in this sense are different from the app delivered by Aldiss et al, which was oriented toward recovery and improved chemotherapy experience [57]. As this difference in delivery time in terms of the different stages in the care pathway demonstrates, the apps described in the included papers address the needs of adolescents with various conditions. These conditions have similarities as well as notable differences. Moreover, there is heterogeneity in the content and delivery of these apps.

There was variation in whether apps were delivered as a stand-alone resource or whether they were used alongside other components of a medical intervention. Burbank et al [39] and Cafazzo et al [55] delivered the app as a stand-alone resource. In the case of the app developed by Aldiss et al, hospital-based nurses were alerted if there was cause for concern; adolescents could also make phone calls to the hospital if they wished [57]. The Frøisland et al study involved a consultation with a health professional midway through the intervention period to give the adolescent an opportunity to discuss and reflect on their use of the app [56].

There was also variation in whether adolescents received guidance in using the app or whether they were solely responsible for navigating the app. Aldiss et al [57] and Burbank et al [39] delivered their apps so that adolescents had sole responsibility in their navigation. Cafazzo et al [55] and Frøisland et al [56] spoke of the adolescents receiving initial training in using the app and the model of mobile phone that would be used in the intervention. The studies also differed in regard to reporting a primary outcome measure; Aldiss et al included six psychometric measures, although they did not identify a primary outcome measure [57]. Cafazzo et al described their primary outcome as an increased frequency of daily blood glucose readings [55]. Frøisland et al simply listed outcomes in the following order: HbA1c, system usability, and knowledge [56]. Burbank et al listed the following outcomes: usage and satisfaction rates, asthma control test, and asthma self-efficacy [39]. One app—developed by Cafazzo et al [55]—was underpinned by the concept of ecological momentary interventions [67], while the other apps were not theoretically driven.

Quality Assessment

Table 2 reports the score on the six individual quality assessment items as scored by the EPHPP tool [34].

Table 2. Study quality assessment for each study.

Quality assessment items	Study and quality assessment score[a]			
	Aldiss [57]	Burbank [39]	Cafazzo [55]	Frøisland [56]
Selection bias	Weak	Weak	Weak	Weak
Study design	Moderate	Weak	Moderate	Weak
Confounders	Weak	Weak	Weak	Weak
Blinding	Moderate	Weak	Weak	Weak
Data collection method	Moderate	Strong	Strong	Weak
Withdrawals and dropouts	Moderate	Strong	Moderate	Strong

[a] Items were scored using the Effective Public Health Practice Project (EPHPP) quality assessment tool.

Inter-rater Reliability

Across the four papers, there was moderate-to-good agreement between the raters on the domains of the EPHPP quality assessment tool, with agreement on 17 of the 24 domains (71% agreement; unweighted κ = .60, 95% CI .33-.87; weighted κ = .71, 95% CI .49-.93). There was excellent agreement on the total domain scores for the four papers (ICC = .83, 95% CI -.17 to .99). The wide confidence intervals reflected the small number of papers assessed.

Effect Sizes

Only two papers provided data for calculating effect sizes and both featured mobile phone apps to help adolescents with type 1 diabetes. Both papers reported small sample sizes; Cafazzo et al [55] reported outcome data for 14-20 participants while Frøisland et al [56] reported data for 11-12 participants, and as these were feasibility studies, they would not have been powered to detect statistical significance. Cafazzo et al reported means and standard deviations before and after their intervention for a wide range of outcome measures [55], but did not report standard deviations for their primary outcome of frequency of blood glucose measurement that showed a significant improvement (*P*=.006), so it was not possible to estimate Cohen's *d*. The change in HbA1c level was numerically small—from 8.8 to 9.2—and nonsignificant, but the effect size was good (*d*=-0.46, *P*=.11). They found no significant changes in dimensions of the Diabetes Self-Care Inventory and the effect sizes varied (adherence *d*=0.11, blood glucose regulation *d*=0, insulin and food regulation *d*=0.12, and emergency preparedness *d*=0). For the only dimension where there was a near significant but sizeable improvement (exercise *d*=0.56, *P*=.07), Cafazzo et al attributed the improvement to a seasonal change from winter to spring as their intervention did not target exercise [55]. While they found no significant improvements in dimensions of the Diabetes Family Responsibility Questionnaire, the effect sizes were medium to large (caregiver's perspective *d*=0.69 and adolescent's perspective *d*=0.72); however, the mean scores only changed by 0.1 point on a 3-point scale. They also found no significant changes in dimensions of the Diabetes Quality-of-Life Instrument for Youth, where the effect sizes were small or small to medium (impact of symptoms *d*=0.36, impact of treatment *d*=-0.07, impact of activities *d*=0.26, parent issues *d*=0.16, worries about diabetes *d*=0.30, and health perception *d*=0.15). Frøisland et al reported means and standard deviations before and after their intervention for HbA1c levels and a knowledge test score [56]. Neither outcome showed a significant difference nor an effect size better than small (HbA1c mean *before* 8.3 [SD 0.9], mean *after* 8.1 [SD 0.9], *d*=0.23, *P*=.38; knowledge test *d*=0.04, *P*=.82).

Patient and Public Involvement in the Included Studies

Since mobile phones and tablets as the new generation of mobile electronic devices are a consumer product with custom software programs called apps, it is perhaps especially important to take into account adolescents' attitudes toward using apps developed to help manage their chronic conditions. While we know mobile apps in general are widely accepted by adolescents living in today's technology-rich environment, we were interested to know whether adolescents had been involved in the development of apps included in this review that were aimed at them and their peers. Three of the four studies we reviewed reported some level of patient involvement in the design, development, and/or evaluation of the app (see Table 3). These three papers reported using qualitative research methods for informing the development of mobile apps for adolescents with cancer [57] and adolescents with type 1 diabetes mellitus [55,56].

Table 3. Involvement of adolescents and their families at different stages of mobile app design, development, and evaluation.

Study	Involvement of adolescents and families at each stage		
	Design	Development	Evaluation
Aldiss et al [57]	Identified the symptoms to be assessed and addressed self-care advice	Tested the symptom report system	Reported perceptions of intervention and reviewed the self-care advice Procedures and technical systems were assessed
Burbank et al [39]	Not reported	Not reported	Not reported
Cafazzo et al [55]	Requirements were obtained through qualitative interviews and focus group sessions	Themes were derived from focus group sessions, which were incorporated into the prototype app	Not reported
Frøisland et al [56]	Adolescents suggested improvements for a future app	Not reported	Technical problems were reported, along with improvements of the existing app

Adolescents were important contributors in the development of ASyMS, an advanced symptom management system utilizing mobile phone technology for adolescents with cancer [57]. The authors reported that during Phase 1 development, adolescents' contributions were essential in identifying which symptoms should be assessed via the ASyMS questionnaire. In Phase 2, adolescents tested the symptom report system. As the software progressed into Phase 3 development, adolescents aged 13-18 years who were receiving chemotherapy were involved in addressing and reviewing the self-care advice. Interestingly, these adolescents suggested improvements to the questionnaire by adding descriptive indicators to clarify specific aspects of the intervention; these had not been identified as areas for improvement by researchers.

Similarly, in Cafazzo et al's pilot study [55], 6 adolescents and their parents informed the design and development of a mobile app to support adolescent self-management of type 1 diabetes.

Focus groups with adolescents and their parents revealed specific requirements that were expressed as four themes: the need for fast, discrete transactions; the role of data collection rather than decision making; overcoming decision-making inertia; and ad hoc information sharing. These opinions were incorporated into the prototype version of the app for testing.

Frøisland et al [56] also involved adolescents in the redesign and evaluation of a mobile visual learning intervention for adolescents with type 1 diabetes, whereby they provided guidance for further development of the mobile app. In addition, adolescents also suggested improvements for the existing app, with requests for additional functionality. The authors reported their intention to implement the adolescents' suggestions when designing the next version of the app.

Clinician Expert Involvement in the Included Studies

Appropriate clinical expertise in the specific condition was evident in all four of the research teams, although it was not always clear whether specific clinical (as opposed to academic) expertise in working with adolescents was represented. Health professional involvement in the modelling stages was reported in three of the four studies [39,55,57]. The study by Aldiss et al was exemplary in view of the meaningful involvement of both adolescents and professionals in the study; in particular, adolescents with specific expertise were involved in the modelling stages of development [57]. Although Cafazzo et al conducted focus groups with health professionals in the modelling stages, no results were presented [55]. In the one study which reported theoretical underpinning [55], involvement of health professionals in translating this into the modelling stage was not reported. In the Frøisland et al paper, it is not clear whether the health professionals had any specific input into the app development [56]. However, the finding that the adolescents' theoretical clinical knowledge was not altered after use of the app demonstrates that the intentions of the app use needs to be linked to a system of assessing its efficacy.

Discussion

Principal Findings

This review clearly demonstrates that despite the large number of health care apps in existence, the evidence base for their benefits for adolescents in personal management of their chronic physical health conditions is limited. The additional contextual challenge is that manufacturers are readily developing apps which are not based on empirical evidence [68]. Studies included in this review were all in the early proof-of-concept phase with few participants, meaning that assumptions about generalizability of the findings to the target population cannot be made. The findings reported are therefore preliminary and would need to be validated by larger-scale research. Comparisons between studies are also difficult as a result of the variability including the short and different follow-up times. While these studies alone do not provide high levels of evidence, they do provide valuable information that paves the way for other studies to inform future development and evaluation of complex app interventions [69].

Adolescent-Specific Issues

As increasing numbers of adolescents with chronic conditions have transferred to adult-centered care, significant declines in treatment adherence have been observed during adolescence and the transition period [70]. Using educational interventions alone to enhance medication adherence is insufficient, but the addition of behavioral elements, such as monitoring and goal setting, rewards, contingency contracting, problem solving, and linking medication taking with established routines, may enhance outcomes [15,71]. That said, the small treatment effects of recent adherence-promoting interventions reflect the need to re-examine their delivery, and the mechanisms of emotional, social, and family processes [72]. Adolescence is arguably a crucial time for the rigorous development, evaluation, and implementation of interventions that promote shared and self-management skills and knowledge, and for the promotion of healthy behaviors [73]. While it is widely recognized that communication technologies are important drivers in adolescent health [4], there are barriers to the use of mobile technologies by adolescents. These include the disparity of access to mobile devices and the potential for habituation, suggesting that the use of information technology (IT) to address health issues may be limited or even harmful to adolescents [74,75].

Participatory Design

Although these studies [39,55-57] support the view that engaging adolescents with chronic conditions has contributed to changes in mobile intervention design, the effects of involvement on accessibility and acceptability (ie, retention and use of mobile apps) was not examined. A consistent finding in this systematic review suggests that adolescents are engaged in helping to design mobile interventions; however, they may not have been actively involved as equal contributing partners in the entire research development and implementation process, as recommended by INVOLVE, the public involvement body funded by the UK National Institute for Health Research. INVOLVE suggests that involvement, engagement, and participation are often linked, and although they are distinct roles, they can indeed complement each other [76]. Examples of patient and public involvement in research include coapplicant responsibility on grants and research projects, involvement in identifying research priorities, membership of project steering groups, and undertaking interviews with research participants [77]. The meaningful involvement, engagement, and participation of adolescents and their families in the entire planning, development, and intervention of mobile apps is likely to contribute to more widely accepted and understood interventions by individuals living with chronic conditions in the future.

Impact of Intervention on Parents, Carers, and Health Care Professionals

Parents and carers also play a significant part in promoting the development of adolescents' personal management skills in chronic conditions [78], but parents may be less confident than adolescents in using technology [79]. Furthermore, given that this is a relatively underdeveloped area of adolescent health services, it can be difficult for those health professionals who are themselves unfamiliar with mobile phone and tablet apps to engage effectively with adolescents via these media [80,81]; future research may consider alternative ways of engaging busy clinicians in research, such as telephone interviews [82].

None of the studies reviewed specifically considered the impact of the technology on parents and their role in the development of adolescents' self-management skills. Similarly, although the mutual benefits of participatory design for both end users and developers were highlighted in the papers, there was little discussion about the implications of app usage on care delivery for health professionals. In one study, there was a mandatory consultation with a health professional to discuss the app [56]. Aldiss et al noted that involvement of professionals during the development process was “the first step in embedding the system into practice” [57] and without this, collaborative and focused clinical care is unlikely to improve despite technical advances and innovations. Expertise in working with adolescents in this area is essential because of the need to consider adolescents' development in the context of chronic conditions, both in terms of clinical care and involvement in research. Training in adolescent health care is not yet universal so it may be difficult for those professionals who have unmet training needs in this area to engage effectively with adolescents via these media [83]. Addressing professional concerns will be important to ensure efficacy of such interventions, for example, Frøisland et al [56] reported that before their study, the participating professionals expressed a fear that SMS text messaging would be overused, but these assumptions proved unfounded.

Limitations

Due to resource limitations, this review excluded mental health conditions and learning disabilities and was only able to consider the three common chronic physical conditions reported in the included studies—asthma, cancer, and diabetes. There is, however, emerging data in the area of mental health care, albeit not specifically targeting adolescents, reporting that the majority of existing apps for mental health care lack scientific evidence about their efficacy [64,84]. In future, therefore, it will be important to reflect on the use of apps for adolescents with other chronic physical, as well as mental health, conditions, as many of the issues facing these adolescents may be similar.

While we used a recognized tool to assess the quality of the four studies, it did not consider factors such as user and health professional involvement in intervention, yet these factors are recognized as being increasingly important in the development and evaluation of complex interventions.

The small evidence base identified by our review emphasizes the need for future high-quality, sophisticated trials in the area of app development for adolescents with chronic conditions, and the total sample size of 46 participants limits generalizability of the findings. The dearth of existing evidence prevented us from commenting on the effectiveness of mobile apps designed to support adolescents' management of their physical conditions, as had been the objective at the outset. This in itself was an important finding and generated stimulating discussion around what the next step should be, from a multi-professional expert audience at an international conference where preliminary findings from this review were reported [85]. A clear recommendation from this work is the need for high-quality RCTs in this field. Given the paucity of papers meeting the criteria for this review, it will be valuable to repeat the review and refine its original objectives in 2-3 years when more evidence is likely to exist.

Our rationale for using the WHO definition of young people aged 10-24 years was based on the specific developmental implications of this age group, which differ considerably from those of younger children and older adults; this meant that papers which were otherwise potentially relevant needed to be excluded. For example, of the 19 full-text papers considered for eligibility, the age range of reported populations for some studies fell outside our predetermined definition. Where we were unable to extract specific data that related to 10-24-year-old participants, either from the papers themselves or by communication with the original authors, we excluded those studies from our analysis. While it is outside the scope of this review to return to the search stage and revise the inclusion and exclusion criteria, it is a consideration for a future review.

Information on adolescents who decline to take part in studies is of major importance in research with this age group, particularly as research nonrespondents have been reported to have poorer health outcomes than those who do respond. In addition, identified barriers to the use of mobile technologies by adolescents, including the disparity of access to mobile devices and the potential for habituation, suggest that the use of IT to address health issues may be limited or even harmful to adolescents [27,28] None of the studies reviewed specifically considered these issues and it will be important for future studies to consider these factors, particularly in studies involving adults where habituation has already been highlighted as a potential limitation of individual apps [75]. Moreover, due to the lack of demographic details in the studies reviewed, comments regarding any gender differences in the use of apps are not possible. This is of interest in light of reports that adolescent females are more likely to access the Internet for health matters than adolescent males [74].

Recommendations from our systematic review emphasize the value of a multidisciplinary team enabling expert clinical and patient involvement in the app design, development, and evaluation stages, as well as the involvement of technological and research personnel. It is also recommended that future trials are based on sound theory and are tested across age groups (where appropriate, and while recognizing the different developmental stages of adolescence), gender, and ethnicity. Future work should also carefully consider which primary and secondary outcomes are important to assess, and the best medium- to long-term follow-up times in context of the longevity and persistence of any behavior change observed. Furthermore, it is anticipated that initially developing a robust, adolescent-friendly app in one condition may have the additional advantage of subsequent adaptability and/or transferability across other conditions.

Conclusions

In conclusion, a key finding of the review is the paucity of evidence-based apps that exist in contrast to the thousands of apps available on the app market that are not evidence-based or user and professional informed. Although we aimed to assess the effectiveness of apps, the dearth of studies meeting our criteria meant that we were unable to be conclusive in this regard. This review provides valuable findings and paves the way for future rigorous development and evaluation of health apps for adolescents with chronic conditions. There remains a need for a phased approach to well-designed trials of mobile phone and tablet apps which resonate with the lives of adolescents, that can be feasibly transferred into real-life settings and which involve adolescents, parents, and health professionals in their design, development, and evaluation. Based on the available evidence, apps may be considered feasible health interventions, but more studies involving larger sample sizes, and with patient and health professional input at all stages, are needed to determine apps' acceptability and effectiveness.

Acknowledgments

This review includes independent research supported by (or funded by) the National Institute for Health Research Biomedical Research Unit Funding Scheme. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health. The authors of this review would also like to thank the University of Manchester Research Institute (UMRI) Pump Priming Programme for providing funding to support the project and we gratefully acknowledge the contributions of Professor Helen Foster and Mrs Sharon Douglas to the project.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Excluded studies and reasons for exclusion.

[PDF File (Adobe PDF File), 5KB - jmir_v17i12e287_app1.pdf]

References

1. Jackson Allen P. The primary care provider and children with chronic conditions. In: Jackson Allen P, Vessey JA, Schapiro N, editors. Primary Care of the Child with a Chronic Condition. 5th edition. St. Louis, MO: Mosby; 2010:3-21.
2. Brooks F, Magnusson J, Klemera E, Spencer N, Morgan A. Health Behaviour in School-aged Children (HBSC): World Health Organization Collaborative Cross National Study. Hatfield, UK: University of Hertfordshire; 2011 Oct. URL: http://www.hbscengland.com/wp-content/uploads/2013/10/HBSC-England-report2011.pdf [accessed 2015-12-03] [WebCite Cache ID 6dVdsev7S]
3. Swallow V. Special issue on children, young people, and families living with long-term conditions. J Pediatr Nurs 2015;30(1):1-3. [doi: 10.1016/j.pedn.2014.10.015] [Medline: 25448470]
4. Sawyer SM, Afifi RA, Bearinger LH, Blakemore S, Dick B, Ezeh AC, et al. Adolescence: A foundation for future health. Lancet 2012 Apr 28;379(9826):1630-1640. [doi: 10.1016/S0140-6736(12)60072-5] [Medline: 22538178]
5. Colver A, Longwell S. New understanding of adolescent brain development: Relevance to transitional healthcare for young people with long term conditions. Arch Dis Child 2013 Nov;98(11):902-907 [FREE Full text] [doi: 10.1136/archdischild-2013-303945] [Medline: 23986559]
6. Ambresin A, Bennett K, Patton GC, Sanci LA, Sawyer SM. Assessment of youth-friendly health care: A systematic review of indicators drawn from young people's perspectives. J Adolesc Health 2013 Jun;52(6):670-681. [doi: 10.1016/j.jadohealth.2012.12.014]
7. Crowley R, Wolfe I, Lock K, McKee M. Improving the transition between paediatric and adult healthcare: A systematic review. Arch Dis Child 2011 Jun;96(6):548-553. [doi: 10.1136/adc.2010.202473] [Medline: 21388969]
8. McDonagh J. Has the gap been bridged yet? Young people in paediatric and adult rheumatology. Rheumatology (Oxford) 2013 Aug;52(8):1349-1351 [FREE Full text] [doi: 10.1093/rheumatology/kes407] [Medline: 23353646]
9. Stinson J, Kohut SA, Spiegel L, White M, Gill N, Colbourne G, et al. A systematic review of transition readiness and transfer satisfaction measures for adolescents with chronic illness. Int J Adolesc Med Health 2014;26(2):159-174. [doi: 10.1515/ijamh-2013-0512] [Medline: 23828488]
10. Department of Health. Improving Children and Young People's Health Outcomes: A System Wide Response. London, UK: Department of Health; 2013 Feb. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/214928/9328-TSO-2900598-DH-SystemWideResponse.pdf [accessed 2015-12-03] [WebCite Cache ID 6dVg0QM2M]
11. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: Qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
12. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: A systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]

13. Gray NJ, McDonagh JE, Harvey K, Prescott J, Roberts R, Shaw KL, et al. Arthriting: Exploring the Relationship Between Identity and Medicines Use, and to Identify the Contribution of Medicines and Pharmacy Services, for the Care of Young People With Arthritis. London, UK: Pharmacy Research UK; 2013 Dec. URL: http://www.pharmacyresearchuk.org/waterway/wp-content/uploads/2013/07/Arthriting-full-report-final-formatted-12dec13.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWgN9jvj]
14. Pai ALH, Ostendorf HM. Treatment adherence in adolescents and young adults affected by chronic illness during the health care transition from pediatric to adult health care: A literature review. Child Health Care 2011;40(1):16-33. [doi: 10.1080/02739615.2011.537934]
15. Dean AJ, Walters J, Hall A. A systematic review of interventions to enhance medication adherence in children and adolescents with chronic illness. Arch Dis Child 2010 Sep;95(9):717-723. [doi: 10.1136/adc.2009.175125] [Medline: 20522463]
16. Pai ALH, McGrady M. Systematic review and meta-analysis of psychological interventions to promote treatment adherence in children, adolescents, and young adults with chronic illness. J Pediatr Psychol 2014 Sep;39(8):918-931 [FREE Full text] [doi: 10.1093/jpepsy/jsu038] [Medline: 24952359]
17. Wu YP, Hommel KA. Using technology to assess and promote adherence to medical regimens in pediatric chronic illness. J Pediatr 2014 Apr;164(4):922-927. [doi: 10.1016/j.jpeds.2013.11.013] [Medline: 24359939]
18. Lorig KR, Holman H. Self-management education: History, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [Medline: 12867348]
19. The Communications Market Report. London, UK: Ofcom; 2014 Aug 07. URL: http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr14/2014_UK_CMR.pdf [accessed 2015-11-27] [WebCite Cache ID 6dLvTHrbe]
20. Wang J, Wang Y, Wei C, Yao NA, Yuan A, Shan Y, et al. Smartphone interventions for long-term health management of chronic diseases: An integrative review. Telemed J E Health 2014 Jun;20(6):570-583. [doi: 10.1089/tmj.2013.0243] [Medline: 24787747]
21. Obiodu V, Obiodu E. An empirical review of the top 500 medical apps in a European Android market. J Mob Technol Med 2012 Dec 22;1(4):22-37. [doi: 10.7309/jmtm.74]
22. Transforming Participation in Health and Care: 'The NHS belongs to us all'. London, UK: Patients and Information Directorate, NHS England; 2013. URL: https://www.england.nhs.uk/wp-content/uploads/2013/09/trans-part-hc-guid1.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWhNDS3T]
23. Aitken M, Gauntlett C. Patient Apps for Improved Healthcare: From Novelty to Mainstream. Parsippany, NJ: IMS Institute for Healthcare Informatics; 2013 Oct. URL: http://www.obroncology.com/imshealth/content/IIHI%20Apps%20report%20231013F_interactive.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWhilbsN]
24. Misra S, Lewis TL, Aungst TD. Medical application use and the need for further research and assessment for clinical practice: Creation and integration of standards for best practice to alleviate poor application design. JAMA Dermatol 2013 Jun;149(6):661-662. [doi: 10.1001/jamadermatol.2013.606] [Medline: 23783150]
25. Department of Child and Adolescent Health and Development. The Second Decade: Improving Adolescent Health and Development. Geneva, Switzerland: World Health Organization; 2001. URL: http://apps.who.int/iris/bitstream/10665/64320/1/WHO_FRH_ADH_98.18_Rev.1.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWi2T05v]
26. UNESCO. What do we mean by "youth"? URL: http://www.unesco.org/new/en/social-and-human-sciences/themes/youth/youth-definition/ [accessed 2015-11-27] [WebCite Cache ID 6dLwDh7Dg]
27. Majeed-Ariss R, Hall AG, McDonagh J, Fallon D, Swallow V. Mobile phone and tablet apps to support young people's management of their physical long-term conditions: A systematic review protocol. JMIR Res Protoc 2015 Apr 07;4(2):e40 [FREE Full text] [doi: 10.2196/resprot.4159] [Medline: 25854293]
28. Green S, Higgins JPT. Cochrane Handbook for Systematic Reviews of Interventions. Chichester, UK: John Wiley & Sons; 2008 Nov. Preparing a Cochrane review URL: https://dhosth.files.wordpress.com/2011/12/cochrane-handbook-for-systematic-reviews-of-interventions.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWj3La1Q]
29. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: Explanation and elaboration. BMJ 2009 Jul 21;339:b2700 [FREE Full text] [Medline: 19622552]
30. Majeed-Ariss R, Hall A, Fallon D, McDonagh J, Swallow V. PROSPERO. York, UK: University of York; 2014. Smartphone and tablet apps for children and young people's management of physical long-term conditions: A systematic review URL: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014015418 [accessed 2015-03-31] [WebCite Cache ID 6XREcWqQY]
31. BBC News. 2003 Mar 03. 3G goes live in the UK URL: http://news.bbc.co.uk/2/hi/technology/2808761.stm [accessed 2015-12-04] [WebCite Cache ID 6dWjTArsr]
32. TMCnet. 2007 Mar 19. GSA announces 100 commercial HSDPA networks worldwide URL: http://www.tmcnet.com/usubmit/2007/03/19/2431009.htm [accessed 2015-12-06] [WebCite Cache ID 6dZwBpoQN]
33. Systematic Reviews: CRD's Guidance for Undertaking Reviews in Health Care. York, UK: Centre for Reviews and Dissemination, University of York; 2009 Jan. URL: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWkzNgfj]

34. Quality Assessment Tool for Quantitative Studies. Hamilton, ON: Effective Public Health Practice Project; 2010. URL: http://www.ephpp.ca/PDF/Quality%20Assessment%20Tool_2010_2.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWlCQoC5]
35. Armijo-Olivo S, Stiles CR, Hagen NA, Biondo PD, Cummings GG. Assessment of study quality for systematic reviews: A comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: Methodological research. J Eval Clin Pract 2012 Feb;18(1):12-18. [doi: 10.1111/j.1365-2753.2010.01516.x] [Medline: 20698919]
36. Altman DG. Practical Statistics for Medical Research. London, UK: Chapman & Hall/CRC; 1990.
37. VassarStats. URL: http://vassarstats.net/ [accessed 2015-11-26] [WebCite Cache ID 6dKYcmydM]
38. Fleiss JL, Cohen J. Educ Psychol Meas. 1973. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability URL: http://www.wpic.pitt.edu/research/biometrics/Publications/Biometrics%20Archives%20PDF/456-Fleiss&Cohen19730001.pdf [accessed 2015-12-04] [WebCite Cache ID 6dWmZv1wR]
39. Burbank A, Lewis SD, Hewes M, Schellhase DE, Rettiganti M, Hall-Barrow J, et al. Mobile-based asthma action plans for adolescents. J Asthma 2015;52(6):583-586. [doi: 10.3109/02770903.2014.995307] [Medline: 25494553]
40. Ryan D, Price D, Musgrave SD, Malhotra S, Lee AJ, Ayansina D, et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: Multicentre randomised controlled trial. BMJ 2012;344:e1756 [FREE Full text] [Medline: 22446569]
41. Stinson JN, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. iCanCope with Pain™: User-centred design of a Web- and mobile-based self-management program for youth with chronic pain based on identified health care needs. Pain Res Manag 2014;19(5):257-265 [FREE Full text] [Medline: 25000507]
42. Kumar N, Khan RA, Kumar P, Sharma VP. E-health: Stipulation of mobile phone technology in adolescent diabetic patient care in pediatric diabetes. In: Proceedings of the the Annual Conference of the International Society for Pediatric and Adolescent Diabetes. 2012 Presented at: Annual Conference of the International Society for Pediatric and Adolescent Diabetes; October 10-13, 2012; Istanbul, Turkey.
43. Scaramuzza A, Bosetti A, Redaelli F, Gazzarri A, Rossi E, Ferrari M, et al. To WhatsApp or not to WhatsApp? What could be done with new social media to manage type 1 diabetes in adolescents. In: Proceedings of the 7th International Conference on Advanced Technologies & Treatments for Diabetes. 2014 Presented at: 7th International Conference on Advanced Technologies & Treatments for Diabetes; February 5-8, 2014; Vienna, Austria.
44. Kovatchev BP. Clinical results from transitional and home trials of outpatient closed-loop control. In: Proceedings of the 7th International Conference on Advanced Technologies & Treatments for Diabetes. 2014 Presented at: 7th International Conference on Advanced Technologies & Treatments for Diabetes; February 5-8, 2014; Vienna, Austria.
45. Stukus D, Phillips G, Farooqui N. Improved education and self-management in children and adolescents with asthma using a personalized smartphone application. In: Proceedings of the Annual Meeting of the American Academy of Allergy, Asthma and Immunology. 2014 Presented at: Annual Meeting of the American Academy of Allergy, Asthma and Immunology; February 28-March 4, 2014; San Diego, CA.
46. Cummings E, Hauser J, Cameron-Tucker H, Fitzpatrick P, Jessup M, Walters EH, et al. Enhancing self-efficacy for self-management in people with cystic fibrosis. Stud Health Technol Inform 2011;169:33-37. [Medline: 21893709]
47. Jacob E, Stinson J, Duran J, Gupta A, Gerla M, Ann LM, et al. Usability testing of a smartphone for accessing a Web-based e-diary for self-monitoring of pain and symptoms in sickle cell disease. J Pediatr Hematol Oncol 2012 Jul;34(5):326-335 [FREE Full text] [doi: 10.1097/MPH.0b013e318257a13c] [Medline: 22627570]
48. Leone JR. Utility of a wireless, handheld monitoring system in the management of hemophilia patients. Comput Inform Nurs 2011 Sep;29(9):521-522. [doi: 10.1097/NCN.0b013e3182066356] [Medline: 21239988]
49. Carroll AE, DiMeglio LA, Stein S, Marrero DG. Contracting and monitoring relationships for adolescents with type 1 diabetes: A pilot study. Diabetes Technol Ther 2011 May;13(5):543-549 [FREE Full text] [doi: 10.1089/dia.2010.0181] [Medline: 21406011]
50. Jessup M, Hauser J, Cameron-Tucker H, Cummings E, Turner P, Blizzard L, et al. Facilitating self-management in adolescents and adults with cystic fibrosis: A pilot study. Pediatr Pulmonol 2011;46(Suppl 34):405.
51. Marciel KK, Saiman L, Quittell LM, Dawkins K, Quittner AL. Cell phone intervention to improve adherence: Cystic fibrosis care team, patient, and parent perspectives. Pediatr Pulmonol 2010 Feb;45(2):157-164 [FREE Full text] [doi: 10.1002/ppul.21164] [Medline: 20054860]
52. McClellan CB, Schatz JC, Puffer E, Sanchez CE, Stancil MT, Roberts CW. Use of handheld wireless technology for a home-based sickle cell pain management protocol. J Pediatr Psychol 2009 Jun;34(5):564-573 [FREE Full text] [doi: 10.1093/jpepsy/jsn121] [Medline: 19029141]
53. McCann L, Maguire R, Miller M, Kearney N. Patients' perceptions and experiences of using a mobile phone-based advanced symptom management system (ASyMS) to monitor and manage chemotherapy related toxicity. Eur J Cancer Care (Engl) 2009 Mar;18(2):156-164. [doi: 10.1111/j.1365-2354.2008.00938.x] [Medline: 19267731]
54. Kumar VS, Wentzell KJ, Mikkelsen T, Pentland A, Laffel LM. The DAILY (Daily Automated Intensive Log for Youth) trial: A wireless, portable system to improve adherence and glycemic control in youth with diabetes. Diabetes Technol Ther 2004 Aug;6(4):445-453. [doi: 10.1089/1520915041705893] [Medline: 15320998]

55. Cafazzo J, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: A pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]
56. Frøisland DH, Arsand E, Skårderud F. Improving diabetes care for young people with type 1 diabetes through visual learning on mobile phones: Mixed-methods study. J Med Internet Res 2012;14(4):e111 [FREE Full text] [doi: 10.2196/jmir.2155] [Medline: 22868871]
57. Aldiss S, Taylor RM, Soanes L, Maguire R, Sage M, Kearney N, et al. Working in collaboration with young people and health professionals. A staged approach to the implementation of a randomised controlled trial. J Res Nurs 2010 Sep 17;16(6):561-576. [doi: 10.1177/1744987110380803]
58. Varni JW, Burwinkle TM, Katz ER, Meeske K, Dickinson P. The PedsQL in pediatric cancer: Reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales, Multidimensional Fatigue Scale, and Cancer Module. Cancer 2002 Apr 1;94(7):2090-2106 [FREE Full text] [Medline: 11932914]
59. Enskär K, Carlsson M, Golsäter M, Hamrin E. Symptom distress and life situation in adolescents with cancer. Cancer Nurs 1997 Feb;20(1):23-33. [Medline: 9033147]
60. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.
61. Gibson F, Aldiss S, Taylor RM, Maguire R, McCann L, Sage M, et al. Utilization of the Medical Research Council evaluation framework in the development of technology for symptom management: The ASyMS-YG Study. Cancer Nurs 2010;33(5):343-352. [doi: 10.1097/NCC.0b013e3181cb4bad] [Medline: 20495448]
62. Gibson F, Aldiss S, Taylor RM, Maguire R, Kearney N. Involving health professionals in the development of an advanced symptom management system for young people: The ASyMS-YG study. Eur J Oncol Nurs 2009 Jul;13(3):187-192. [doi: 10.1016/j.ejon.2009.03.004] [Medline: 19409843]
63. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Erlbaum Associates; 1988.
64. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: A systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
65. Grissom RJ, Kim JJ. Effect Sizes for Research: A Broad Practical Approach. Mahwah, NJ: Erlbaum Associates; 2005.
66. Daniel T, Kostic B. RStats Institute, Missouri State University. 2015. RStats effect size calculator URL: http://www.missouristate.edu/rstats/Tables-and-Calculators.htm [accessed 2015-12-04] [WebCite Cache ID 6dWsvm0DL]
67. Heron KE, Smyth JM. Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
68. Buijink AW, Visser BJ, Marshall L. Medical apps for smartphones: Lack of evidence undermines quality and safety. Evid Based Med 2013 Jun;18(3):90-92. [doi: 10.1136/eb-2012-100885] [Medline: 22923708]
69. Craig C, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: The new Medical Research Council guidance. BMJ 2008;337:a1655 [FREE Full text] [Medline: 18824488]
70. Pai ALH, Ostendorf HM. Treatment adherence in adolescents and young adults affected by chronic illness during the health care transition from pediatric to adult health care: A literature review. Child Health Care 2011 Jan;40(1):16-33. [doi: 10.1080/02739615.2011.537934]
71. Santer M, Ring N, Yardley L, Geraghty AW, Wyke S. Treatment non-adherence in pediatric long-term medical conditions: Systematic review and synthesis of qualitative studies of caregivers' views. BMC Pediatr 2014;14:63 [FREE Full text] [doi: 10.1186/1471-2431-14-63] [Medline: 24593304]
72. Hood KK, Rohan JM, Peterson CM, Drotar D. Interventions with adherence-promoting components in pediatric type 1 diabetes: Meta-analysis of their impact on glycemic control. Diabetes Care 2010 Jul;33(7):1658-1664 [FREE Full text] [doi: 10.2337/dc09-2268] [Medline: 20587726]
73. Heron KE, Smyth JM. Ecological momentary interventions: Incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
74. Suris J, Akré C, Berchtold A, Bélanger RE, Michaud P. Chronically connected? Internet use among adolescents with chronic conditions. J Adolesc Health 2010 Feb;46(2):200-202. [doi: 10.1016/j.jadohealth.2009.07.008] [Medline: 20113927]
75. Suris JC, Akre C, Piguet C, Ambresin AE, Zimmermann G, Berchtold A. Is Internet use unhealthy? A cross-sectional study of adolescent Internet overuse. Swiss Med Wkly 2014;144:w14061 [FREE Full text] [doi: 10.4414/smw.2014.14061] [Medline: 25474244]
76. INVOLVE. London, UK: NIHR What is public involvement in research? URL: http://www.invo.org.uk/find-out-more/what-is-public-involvement-in-research-2/ [accessed 2015-11-27] [WebCite Cache ID 6dMBZhRlt]
77. McDonagh JE, Bateman B. 'Nothing about us without us': Considerations for research involving young people. Arch Dis Child Educ Pract Ed 2012 Apr;97(2):55-60. [doi: 10.1136/adc.2010.197947] [Medline: 21803922]

78. Nightingale R, Sinha MD, Swallow V. Using focused ethnography in paediatric settings to explore professionals' and parents' attitudes towards expertise in managing chronic kidney disease stage 3-5. BMC Health Serv Res 2014;14:403 [FREE Full text] [doi: 10.1186/1472-6963-14-403] [Medline: 25234741]
79. Swallow V. An exploration of mothers' and fathers' views of their identities in chronic-kidney-disease management: Parents as students? J Clin Nurs 2008 Dec;17(23):3177-3186. [doi: 10.1111/j.1365-2702.2008.02642.x] [Medline: 19012785]
80. Hardoff D, Danziger Y, Reisler G, Stoffman N, Ziv A. Minding the gap: Training in adolescent medicine when formal training programmes are not available. Arch Dis Child Educ Pract Ed 2009 Oct;94(5):157-160. [doi: 10.1136/adc.2008.155762] [Medline: 19770496]
81. Swallow V, Clarke CL, Iles S, Harden J. Work based, lifelong learning through professional portfolios: Challenge or reward? Pharm Educ 2006;6(2):77-89.
82. Marr S, Steele K, Swallow V, Craggs S, Procter S, Newton J, et al. Mapping the range and scope of emergency nurse practitioner services in the Northern and Yorkshire Region: A telephone survey. Emerg Med J 2003 Sep;20(5):414-417 [FREE Full text] [Medline: 12954677]
83. Brown T. Are you a digital native or digital immigrant: Being client centred in the digital era. Br J Occup Ther 2011 Jul;74(7):313.
84. Farrer L, Gulliver A, Chan JK, Batterham PJ, Reynolds J, Calear A, et al. Technology-based interventions for mental health in tertiary students: Systematic review. J Med Internet Res 2013;15(5):e101 [FREE Full text] [doi: 10.2196/jmir.2639] [Medline: 23711740]
85. Majeed-Ariss R, Fallon D, Hall A, McDonagh J, Swallow V. Health theory in mobile technology apps supporting young people's long-term condition/s management: A systematic review [in press]. Bull Eur Health Psychol Soc 2015 in press.

Abbreviations

AAP: asthma action plan
ASyMS: Advanced Symptom Management System
ASyMS-YG: Advanced Symptom Management System for young people
CINAHL: Cumulative Index to Nursing and Allied Health Literature
Cond.: condition
EPHPP: Effective Public Health Practice Project
F: female
HbA1c: hemoglobin A1c (glycated hemoglobin)
ICC: intraclass correlation coefficient
ID: intervention duration
IMS: Institute of Medical Science
IMS Health: Institute of Medical Science Institute of Healthcare Informatics
IT: information technology
LSS-A: Life Situation Scale for Adolescents
N/A: not applicable
NHS: National Health Service
NIHR: National Institute for Health Research
PICOS: population or participant, intervention or indicator, comparator or control, outcome, and study design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO: international prospective register of systematic reviews
QOL: quality of life
RCT: randomized controlled trial
SMS: short message service
SS: sample size
STAI: State-Trait Anxiety Inventory
UMRI: University of Manchester Research Institute
WCDMA: wideband code-division multiple access
WHO: World Health Organization

Edited by G Eysenbach; submitted 05.09.15; peer-reviewed by Y Wu, J Wang; comments to author 24.09.15; revised version received 23.10.15; accepted 15.11.15; published 23.12.15.

Please cite as:
Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, McDonagh JE, Stones SR, Thomson W, Swallow V
Apps and Adolescents: A Systematic Review of Adolescents' Use of Mobile Phone and Tablet Apps That Support Personal Management of Their Chronic or Long-Term Physical Conditions
J Med Internet Res 2015;17(12):e287
URL: http://www.jmir.org/2015/12/e287/
doi:10.2196/jmir.5043
PMID:26701961

Original Paper

Preventing Alcohol Abuse Through Social Networking Sites: A First Assessment of a Two-Year Ecological Approach

Valentin Flaudias[1,2], PhD; Ingrid de Chazeron[1,2], PhD; Oulmann Zerhouni[3], PhD; Jordane Boudesseul[3]; Laurent Begue[3]; Renaud Bouthier[4]; Christel Lévrier[4]; Pierre Michel Llorca[1,2]; Georges Brousse[1,2]

[1]CHU Clermont-Ferrand, Pôle Psychiatrie B, Clermont-Ferrand, France
[2]Clermont Université, EA NPsy-Sydo, Université d'Auvergne, Clermont-Ferrand, France
[3]LIPPC2S, Université Grenoble Alpes, F-38000 Grenoble, France
[4]Association Avenir Santé, Lyon, France

Corresponding Author:
Valentin Flaudias, PhD
CHU Clermont-Ferrand
Pôle Psychiatrie B
58 Rue Montalembert
Clermont-Ferrand, F-63000
France
Phone: 33 473752072
Fax: 33 473752074
Email: vflaudias@chu-clermontferrand.fr

Abstract

Background: Prevention strategies to reduce alcohol use/consumption among young people are crucial to reducing alcohol-related deaths and preventing disease. This paper focuses on the effectiveness of a social networking site (SNS) alcohol prevention program targeted toward young people.

Objective: We hypothesized that the program would diminish the relation made by participants between alcohol and festive moments, and would result in a reduction of their declared consumption of alcohol at festive moments during the program. We also explored the interaction with the prevention program that was the most efficient.

Methods: The prevention program took the form of 3 lotteries over 2 years. The participants periodically received prevention messages, particularly on alcohol and festive moments (eg, videos on Facebook and short message service [SMS] text messages on their mobile phones). For the 3 periods, the participants had to answer questions exploring the level of their belief that alcohol consumption and festive moments are highly associated. A control group that did not participate in the prevention program was asked the same questions over the same number of days for the first 2 periods. During the second period, the participants were asked to answer questions about their alcohol consumption during parties. During the third period, we explored the interaction with the prevention program on the reduction of their belief that alcohol consumption and festive moments are associated.

Results: A total of 651 participants (age: mean 22.24, SD 4.10 years; women: n=430) during the first period, 301 participants (age: mean 21.27, SD 3.07 years; women n=199) during the second period, and 305 (age: mean 22.41, SD 4.65 years; women: n=190) during the third period correctly completed the survey. For the control group, 69 students completed the survey during the first period (age: mean 18.93, SD 1.14 years; women: n=59) and 50 during the second (age: mean 20.78, SD 1.94 years; women: n=45). We observed a significant reduction in the association of alcohol with festive moments in the participants over the 2 years (period 1: $z=–4.80$, $P<.001$; period 2: $z=–2.11$, $P=.04$; period 3: $z=–2.30$; $P=.02$), but not in the controls. We also observed a reduction in the number of glasses consumed during festive moments for the participants ($z=–2.36$, $P=.02$), but not for the controls during the second period. The third period showed that only the number of days since registration in the program had an impact on the reduction of the association of festive moments and alcohol consumption ($t_{21}=3.186$, $P=.005$).

Conclusions: The findings of this study suggest that the SNS prevention program is promising in preventing the association of alcohol with festive moments and, more generally, in impacting social norms.

(J Med Internet Res 2015;17(12):e278) doi:10.2196/jmir.4233

KEYWORDS
social networking; primary prevention; alcohol consumption; students

Introduction

Alcohol is the most harmful substance in the United Kingdom [1] and it is the third leading preventable cause of death in the United States [2]. Globally, approximately 76.3 million people have been diagnosed with alcohol use disorders. In most Western countries, alcohol use commonly begins during adolescence [3]. According to a French national survey (Enquête sur la Santé et les Consommations lors de l'Appel de Préparation À la Défense [ESCAPAD]), in 2011, 10% of 17-year-olds participated in regular drinking (at least 10 drinks during a month); 28% reported repeated drunkenness (having been drunk at least 3 times a year), 10.5% were drunk 10 or more times a year, and 53.2% reported having drunk more than 5 glasses of alcohol during one event within the last 30 days [4]. This survey shows an increase in binge drinking in this population. A recent review [5] defined *binge drinking* as “a pattern of drinking alcohol that brings blood alcohol concentration to 0.08 g/dL or above (≥5 for men or ≥4 for women in 2 hours) on more than one occasion within the past 6 months.” This massive consumption during a short period could have tragic short-term effects, such as accidents, violence, or ethyl coma [6]. Recent studies have also shown that this type of consumption seems to have a long-term impact on spatial working memory [7] and other neurocognitive functions [8]. In this paper, the focus is on festive moments, or event occasions. In their review, Mallett et al [9] showed that there are many events during student life. “Festive moments” include all these events: parties, celebrations, sport events, holidays, school breaks, and personally relevant events (see also [10]).

To better understand how the consumption of alcohol during festive moments is integrated, it is important to know how the representation of festive moments is communicated. The media has an important impact on young people about social norms regarding alcohol consumption [11], including that alcohol consumption is mandatory to enjoy a party [12].

Recently, the most important type of media for young people is the Internet, especially social network sites (SNSs). A total of 82% of French people younger than age 25 years use Facebook [13]. Facebook is currently the most popular SNS in the world, topping 1 billion active users with 580 million who engage with the site daily [13-15]. This platform represents an important strategic issue for communicating with young people.

One of the strengths of Facebook is that registered people become vectors of communication for messages posted on a page. Indeed, “liking” a page and commenting on the news will result in the commentary being displayed publicly so that their friends will, in turn, receive this information. Alcohol producers use various strategies on social networks to promote alcohol consumption (eg, interactive games, contests, videos sent to minors, Facebook pages). Nicholls [16] showed that social norms could be influenced by alcohol marketing strategies on SNSs and that “traditional notions of celebration play a key role.” Ridout et al [17] observed a normalization of alcohol consumption, particularly in the search for an acceptable social identity, because of the importance of the “drinker” status on SNS. Online interactions contribute to the normalization of youth consumption of alcohol [18]. A recent example regards a beer seller, who launched a lottery to win beer-related prizes. The participants had to register on their Facebook pages and then answer 3 questions about the company. These questions, of course, required the participant to explore the page or to search the Internet for answers. After having responded and participated in the lottery, this information was displayed on the Facebook pages of the participants, allowing their friends to access it as well. At the time this paper was written, this French page had 229,934 fans.

A recent study in the United States showed that teens who use SNSs at least once per day (70% of those aged 12-17 years use them for an average of 23 minutes per day) were 5 times more likely to use tobacco (10% vs 2%), 3 times more likely to consume alcohol (26% vs 9%), twice as likely to use marijuana (13% vs 7%), and almost 4 times as likely to be exposed to images of young people smoking, drinking, or using drugs [19].

Despite the success of these marketing strategies, few prevention programs have used SNSs. Capurro et al [15] recently reviewed 58 articles related to public health research on Internet sites. However, many of these studies focused on users and usage of SNSs rather than the effectiveness of this type of information on prevention.

The main objective of this paper was to evaluate a SNS-based prevention program [20] and show the results of its effectiveness in changing the representation of festive moments and, more particularly, in altering the high association of alcohol consumption and these moments. For this purpose, we thought that the participants in a 2-year existing SNS-based prevention program would show a weaker association between festive moments and alcohol consumption at the end of the program compared to the beginning. More importantly, we hypothesized that this difference would not be significant for a similar sample questioned at the same time, but who did not register for this program. Finally, our evaluation took into account the behavior (eg, activities) on the program’s SNS page to better understand the reduction in the association between alcohol and festive moments.

Methods

Description of the Prevention Program

This program aimed to reduce extreme alcohol consumption in a festive context among youth. To do this, we used 2 types of communication associated with festive contexts: the Internet through social networks (Facebook) and mobile phone services by sending short message service (SMS) text messages.

Concerning the Internet and, particularly, Facebook, the aim was to regularly spread prevention campaigns with a dedicated Facebook page entitled “Auvernight.” This page contained mostly videos, but also posters and slogans from national programs (eg, the program SAM for drivers, which was aimed toward choosing a sober driver in a group while they are partying) to challenge young people to change their festive habits. Over the 2-year period, approximately 43% (16/37) of these preventive messages directly concerned alcohol; others concerned drug consumption, road accidents, and sexually transmitted diseases. To make the page more attractive and representative of the festive world, we also posted a selection of local festive activities. To bolster interest in this page, we introduced a lottery each academic year (3 lotteries during the 2 years of the program) with attractive prizes (eg, a trip or a computer costing €3500 or approximately US $4000).

Because it is known that alcohol-related cognitions in long-term memory have a strong influence on drinking behavior (eg, expectancies toward alcohol [21,22]) and that alcohol-related contextual cues are likely to activate behavioral schemes associated with these expectancies [23], SMS text messages were sent as close in time as possible to key moments for festive events (eg, when approaching the weekend, at the beginning of school holidays) to ensure a maximal impact on drinking behavior while partying. The messages reminded the participants of a number of tips to reduce the risks potentially associated with alcohol and the negative effects of massive alcohol consumption. Of these SMS text messages, 71% (12/17) specifically concerned alcohol and 29% (5/17) were about other drug consumption. This approach was complementary to the approach used on the Internet in the sense that it was more situational (eg, sending a message during a festive moment such as a New Year celebration).

Regarding this type of communication, it is important to have a long-term perspective to increase the knowledge of this page in the target population. This is why a multiple year program was necessary. To maintain interest, 3 lotteries were organized during the 2 years of the program. Each lottery lasted 3 months. The first took place between February and June 2013, the second between April and June 2014, and the third between October and December 2014. The population could be different for each lottery, so we explored the results for each lottery period independently.

Participants and Procedure

The participants in this program were recruited using emails sent to their personal mailboxes from a listing of the “Avenir Santé” association for their other prevention activities or by flyers distributed near the amphitheater or the library at the University of Clermont-Ferrand (France). This recruitment was conducted for each period. They had to complete an online registration. During this registration, participants answered questions to assess their views on alcohol consumption during festive moments. Three months later, they completed the questionnaires assessing the same questions. To be included in the lottery, they had to be registered on the Facebook page and they were informed that they would have to conserve the SMS text messages on their mobile phones to have an advantage because they would be asked questions about their content. This was done to ensure that they would read the SMS text messages. These participants composed the All Facebookers group.

In the first 2 periods, we chose to include a control group that was composed of students. For the first period, psychology undergraduates were recruited near the University of Clermont-Ferrand using the same methods and places as those used for the participants in the program. They completed the same questionnaire as the All Facebookers group at the same time. For the second period, psychology undergraduate students at the University Grenoble-Alpes (France), who were not registered in the lottery, were involved in the evaluation. They were recruited by email to complete the survey at the beginning and the end of the program. These students composed the control group. We had no control group for the third period.

To ensure that the control group was comparable with the All Facebookers group, we chose to pair students who participated in the prevention program (and so who were members of the All Facebookers group) by sex and age with students of the control group. This was the Paired Facebookers group.

Therefore, we had 3 groups: one group for all participants on Facebook (All Facebookers), students not registered for the prevention program (control), and paired participants/students (Paired Facebookers). The first period of this program took place between February and June 2013 (period 1), the second period between April and June 2014 (period 2), and the third period between October and December 2014 (period 3).

Measurements

Period 1

After demographic data were collected (age and sex), the participants were asked, “On a scale from 0 to 10 (0=not at all, 10=absolutely), how much do you think that alcohol is necessary to have a successful party?” This question was asked online at the beginning and at the end of the program for the Facebookers and on paper for the control group.

Period 2

The 4 questions concerning party representations were “On a scale from 0 to 10 (0=not at all, 10=absolutely):

1. How much do you think that alcohol is necessary to have a successful party?
2. How much do you think that alcohol improves the mood of a party?
3. How much do you think that alcohol may lessen the interest of a party?
4. How much do you think a party is more successful without alcohol?”

To finish, we assessed their alcohol consumption for one party with the following question: “How many glasses do you drink when you are at a party?” The order in which the questions were presented was randomized for each participant. All participants and control group members answered online at the beginning and at the end of the program.

Period 3

The same questions as those asked during period 2 were used at the beginning and at the end of period 3. We added a memory question regarding the SMS text message. For this, we asked the participants to choose between a set of 10 propositions, of which 5 included correct content of the SMS text message that they had received. At the end of the program, the number of likes on the page was counted to measure the activity on the page.

Statistical Analysis

The data were analyzed using SPSS version 19. The comparison between the pre- and posttests (at the beginning and at the end of the period) was computed with a nonparametric Wilcoxon test for the question associated with alcohol and festive moments, and for each sample (All Facebookers, Paired Facebookers, and control). Then, the difference between the pre- and posttests for this association in the Paired Facebookers and the control groups were compared with a Mann-Whitney nonparametric test. When significance was observed, a Cohen's *d* was calculated to assess the effect size. The differences concerning the association of alcohol between these 2 populations at the beginning of the program were compared using a nonparametric Wilcoxon test. The same analysis was conducted for the number of glasses consumed at one festive moment for the pre- and posttests.

Concerning periods 2 and 3, Cronbach alpha analyses were completed to assess the reliability of all questions before a mean was calculated for the 4 questions regarding the association of alcohol with festive moments.

Finally, for the third period, a linear regression analysis was computed on the reduction of the association of alcohol with festive moments and the number of correct SMS text messages remembered, the number of days following registration in the program, and the number of Facebook likes as factors to better understand the reduction.

Results

Description of the Population and Reliability of Items

At the beginning of the first period, we had 866 participants. Three months later, at the end of the program, 651 had completed the questionnaires assessing the same questions (see Table 1). A total of 69 participants were in the control group after one was excluded because of missing data (ie, the questionnaire was not filled out completely); mean age was 18.93, SD 1.14 years and 59 (86%) were women. There were 651 in the All Facebookers group (age: mean 22.24, SD 4.10 years; women: n=430, 66.1%) and 69 in the Paired Facebookers group (age: mean 19.75, SD 0.94 years; women: n=59, 86%).

During the second period, 498 persons participated at the beginning of the program, 424 appropriately answered the 4 questions, and 301 of these correctly completed the 2 questionnaires at the beginning and at the end of the program (14.7%, 52/353 were excluded because they answered inconsistently on negative version questions). There were 21 participants common to the first period. Analyses were performed for the 2 populations: the 301 total participants and the 280 new participants. Thus, 301 participants composed the All Facebookers group (age: mean 21.27, SD 1.94 years; women: n=199, 66.1%) or 280 when we only observed new participants (age: mean 21.24, SD 3.12 years; women: n=185, 66.1%), 50 controls (age: mean 20.78, SD 1.94 years; women: n=45, 90%) and 50 Paired Facebookers (age: mean 20.80, SD 2.00 years; women: n=45, 90%).

In the third period, 452 persons participated at the beginning of the program and 305 appropriately answered the 4 questions and completed the 2 questionnaires at the beginning and at the end of the program (1 was excluded because he answered inconsistently on negative version questions). There were 25 participants who were also involved in the first or the second period. Analyses were performed on the 2 populations: the 305 total participants (age: mean 22.41, SD 4.65 years; women: n=190, 62.3%) and the 280 new participants (age: mean 22.30, SD 4.74 years; women: n=173, 61.8%).

Concerning the reliability of the items (see Table 2), all Cronbach alphas were greater than .70 for the second period and .80 for the third period, which could be considered "acceptable" [24].

Association of Alcohol Consumption with Festive Moments

For the All Facebookers group, we observed a significant reduction in the association of festive moments with alcohol consumption between the beginning and the end of the program for the 3 periods of the evaluation (period 1: $z=–4.80$, $P<.001$; period 2: $z=–2.11$, $P=.04$; period 3: $z=–2.3$, $P=.02$) (see Table 1). Similar results were observed with only new participants for the second and third periods.

For the control participants, we observed that there was no significant difference between the association of alcohol with festive moments at the beginning and at the end of the program for the first and second periods (period 1: $z=–0.35$, $P=.73$; period 2: $z=–0.73$, $P=.47$).

Finally, when we compared the Paired Facebookers and the control groups concerning the differences between pre- and posttest results, we found a difference for the association of alcohol with festive moments only for the first period ($z=–2.24$, $P=.02$).

Concerning the pretest, we observed a difference between our controls and the Paired Facebookers for the first period ($z=–2.13$, $P=.03$), but not for the second ($z=–0.33$, $P=.74$).

Table 1. Comparison between pre- and posttest scores for association of alcohol and festive moments for each group during the 3 test periods and the mean number of glasses of alcohol consumed during one festive moment for periods 2 and 3.

Type of participants and characteristics of analyses	Period 1	Period 2				Period 3			
		All participants		New participants only		All participants		New participants only	
	Alcohol association	Alcohol association	Alcohol consumed during one festive moment	Alcohol association	Alcohol consumed during one festive moment	Alcohol association	Alcohol consumed during one festive moment	Alcohol association	Alcohol consumed during one festive moment
All Facebookers									
n	651	301	301	280	280	305	305	280	280
Pretest, mean (SD)	3.41 (2.72)	4.23 (2.29)	2.68 (1.64)	4.22 (2.26)	2.62 (1.62)	4.34 (2.20)	2.77 (1.60)	4.35 (2.20)	2.80 (1.62)
Posttest, mean (SD)	3.06 (2.45)	4.06 (2.04)	2.53 (1.51)	4.02 (2.04)	2.51 (1.52)	4.15 (2.18)	2.72 (1.57)	4.13 (2.19)	2.72 (1.59)
z	–4.80	–2.11	–2.36	–2.03	–2.09	–2.30	–0.92	–2.36	–1.37
P	<.001	.04	.02	.04	.04	.02	.36	.02	.17
Cohen’s *d*	0.14	.08	.10	.09	.07	.09	—	.10	—
Control									
n	69	50	50	—	—	—	—	—	—
Pretest, mean (SD)	2.83 (2.82)	4.19 (2.51)	2.58 (1.77)	—	—	—	—	—	—
Posttest, mean (SD)	2.93 (2.58)	4.04 (2.54)	2.62 (1.71)	—	—	—	—	—	—
z	–0.35	–0.73	–0.33	—	—	—	—	—	—
P	.73	.47	.74	—	—	—	—	—	—
Cohen’s *d*	—	—	—	—	—	—	—	—	—
Paired Facebookers									
n	69	50	50	—	—	—	—	—	—
Pretest, mean (SD)	1.86 (2.25)	4.27 (2.03)	2.9 (1.47)	—	—	—	—	—	—
Posttest, mean (SD)	1.28 (1.81)	4.32 (1.73)	2.58 (1.43)	—	—	—	—	—	—
z	–2.86	–.50	–2.38	—	—	—	—	—	—
P	.004	.62	.018	—	—	—	—	—	—
Cohen’s *d*	.28	NS	.22	—	—	—	—	—	—
Control vs Paired Facebookers									
z	–2.24	.47	–1.9	—	—	—	—	—	—
P	.025	.64	.057	—	—	—	—	—	—
Cohen’s *d*	–0.34	NS	–0.38	—	—	—	—	—	—

Table 2. Cronbach alphas for the 4 questions on alcohol and its association with festive moments for pre- and posttests and for participant group in the second and third periods.

Type of participants	Period 2, Cronbach α		Period 3, Cronbach α	
	Pretest	Posttest	Pretest	Posttest
All Facebookers	.804	.754	.814	.822
Control	.882	.890	—	—
Paired Facebookers	.772	.708	—	—

Effects on the Number of Glasses Consumed at Festive Moments

For the All Facebookers group, we observed a significant reduction in the number of glasses consumed at festive moments in the second period (*z*=–2.36, *P*=.02), but not the third (*z*=–0.92, *P*=.36).

For the control group, during the second period, we observed that there was no difference between the number of glasses at the beginning and at the end of the program (*z*=–0.33, *P*=.74). Concerning the Paired Facebookers, we observed a reduction between the beginning and the end of the period for the number of glasses consumed at festive moments (*z*=–2.38, *P*=.02). Concerning only the beginning of the period, we observed no difference between our controls and the Paired Facebookers in the number of glasses at festive moments (*z*=–0.79, *P*=.43).

Finally, when we compared the Paired Facebookers and the control group concerning the difference between the beginning and the end of period 2, we observed a difference in the number of glasses per festive moment indicating that the Paired Facebookers had a reduction in the declared number of glasses per festive moment compared to the control participants, but this did not reach statistical significance (*z*=–1.90, *P*=.06).

Effect of the Prevention Program on the Reduction of the Association Between Alcohol Consumption and Festive Moments

For the third period, all participants could recall a mean 4.05 (SD 1.65) SMS text messages, made a mean 0.12 (SD 0.50) likes, and the number of days since registration was a mean 49.48 (SD 7.31) days. The linear regression analyses for the third period (R^2=–.008; $F_{4,303}$=0.366, *P*=.83) did not show any effect of our variables of interest on the reduction of the association between alcohol and festive moments. Viewing the low number of participants with likes on the Facebook page in our population, we decided to explore only participants who had at least one participation on the Facebook page, for example, one like or a comment (n=22; likes: mean 1.68, SD 0.89). The linear regression analysis on this population (R^2=.253; $F_{4,21}$=2.779, *P*=.06) (see Table 3) showed that only the number of days since registration in the program had a significant impact (t_{21}=3.186, *P*=.005) on the difference concerning the association between festive moments and alcohol consumption.

Table 3. Linear regression analyses predicting reduction in association between alcohol and festive moments for participants who have participated at least once, controlling for age, number of Facebook likes, number of SMS text message recalls, and number of days since registration.

Predictors of reduction in association of alcohol and festive moments	β standardized	t_{21}	*P*
Age	.072	0.354	.73
Number of likes	.043	0.207	.84
Number of correct SMS text message recalls	.019	0.086	.93
Number of days since registration	.624	3.186	.005

Discussion

We examined the effectiveness of an SNS-based prevention program on social representations concerning alcohol consumption and party habits. This program was conducted through the Internet over 2 years with 651 participants during the first period, 301 during the second, and 305 during the last period. A total of 1011 different persons participated in this program and this evaluation. We had 69 students as a control group during the first period and 50 during the second period. The results showed a reduction in the link between alcohol and partying for our target population for the 3 periods. This result was not found for our control group. Interestingly, we observed that the declared number of glasses of alcohol consumed at festive moments diminished between the beginning and the end of the program for participants in the second period; however, this was not the case for our control group (or for our third period). Our results failed to show what factors influence this reduction for the population. But, when we explored only participants who had participated at least once, we showed that the reduction in the link between alcohol and festive moments is only influenced by the number of days since registration and not by the age, number of correct SMS text message recalls, or the number of likes on the page.

The results of this evaluation support the recent interest in Web-based programs for health policies [15] and their efficacy. They highlight the important role of this new way of communication for a medical approach. This type of support has many advantages for health promotion. First, our message could be spread to a large audience and not only population at school (pupils). This last type of population is intensively targeted by prevention programs because of the ease of access.

The major interest of SNSs is that it could target other populations. There is no geographical limitation and rural environments could easily have access to this message. Another advantage is its accessibility at all times because it is the participant who can choose when they want to see the message. Moreover, we could adapt the message to a specific population.

Similar health programs via SNSs should use the specificity of this means of communication: the viral nature of information, the engagement of the participant to have an active action on messages, and all messages must be short (1-4 minutes for a video is recommended [25]).

Despite these interests, players in prevention must be very reactive to be effective because there are many changes in this domain and the habits of the young people of today will not remain constant for the long term. This was the case when young people left MSN Messenger (a social network platform) for Facebook during the early 2000s. For example, there is a new social network that is now used by young people: Twitter [26]. Communication on this social network is not exactly the same as on Facebook because users can only type very short messages (140 characters); thus, it is used differently.

There are some limitations to consider when interpreting the results of our study, which are essentially due to the ecological nature of this study. First, we could question our control population, who were mostly students and who may not be very similar to our sample. Their participation was completely uncompensated; thus, they did not have the same motivation as our experimental group. Another limitation is that we assessed the reduction in the association between alcohol and festive moments only for participants who successfully completed the 2 questionnaires. Despite our efforts, it must be noted that it is possible that this type of program will affect only participants who are sensitive to prevention messages. In all cases, we observed a positive impact on approximately 1100 persons.

Another limitation is that our program is not for a specific festive context. Indeed, Mallett et al [9] reviewed some experiences that showed that there are different high-risk events that are associated with alcohol consequences and probably with different consumption patterns. Futures studies should take these differences into account for a more efficient message.

Moreover, our measurements were based on self-reports. We cannot rule out that our results could be due to demand bias from the participants. Because most of our participants participated in a lottery, one cannot exclude the possibility that the participants consciously biased their responses to “please” the experimenter, particularly because there were high incentives for winning and alcohol consumption is especially sensitive to normative pressure [27]. However, Becona [28] found a close relationship between carbon monoxide levels and self-reported smoking rates in a smoking population. Moreover, Cherpitel [29] showed that self-reports of alcohol consumption measured in patients admitted to the emergency room were comparable to blood alcohol concentration. These 2 measures both predict behaviors in a similar manner, such as relating to alcohol-related violence [30; more details on Auvernight are available in 31]. Therefore, because confidentiality and anonymity are assured, we can be confident in our results.

Nevertheless, this ecological and powerful design allows us to believe that a long-term experiment to attest to the deep efficiency of this program could be implemented in the near future. Finally, this evaluation was focused only on alcohol, which was the major objective of this program of prevention, but we also used other messages of prevention (eg, cannabis, sexually transmitted infections). In future studies, it would be interesting to observe the impact of these messages on associated behavior and to observe the impact of the number and the quality of the messages on it. Our assessment failed to show what interaction modulates the association between festive moments and alcohol. Future studies need to explore this, probably utilizing a randomized design.

To conclude, our 2-year study exploring the efficacy of a prevention program on SNSs has shown encouraging results; we observed a reduction in the association of alcohol with festive moments and a reduction in the declared consumption of alcohol while partying. These results show that SNSs could be an interesting type of communication for promoting health and, more particularly, for impacting the social norms associated with alcohol consumption. In the future, the evaluation of the long-term impacts and the exploration of exactly which messages were the most efficient may be of interest.

Acknowledgments

This work was supported by grants from the Mission interministérielle de lutte contre les drogues et les conduites addictives (MILDECA), the Agence régionale de santé (ARS) Auvergne, the Institut national de prevention et d’éducation pour la santé (INPES) and the Direction générale de la santé (DGS). The authors thank the UFR de Psychologie, Université Blaise Pascal Clermont-Ferrand and all the undergraduate students who agreed to participate in our study.

Conflicts of Interest

None declared.

References

1. Nutt DJ, King LA, Phillips LD, Independent Scientific Committee on Drugs. Drug harms in the UK: a multicriteria decision analysis. Lancet 2010 Nov 6;376(9752):1558-1565. [doi: 10.1016/S0140-6736(10)61462-6] [Medline: 21036393]
2. McGinnis JM, Foege WH. Actual causes of death in the United States. JAMA 1993 Nov 10;270(18):2207-2212. [Medline: 8411605]

3. Oei T, Hasking P. Alcohol use disorders. In: Miller P, editor. Comprehension Addictive Behaviors and Disorders, Volume 1: Principles of Addiction. San Diego: Academic Press; 2013:647-655.
4. Observatoire français des drogues et des toxicomanies. Drogues et addictions, donnéees essentielles. Paris: OFDT; 2013.
5. Courtney KE, Polich J. Binge drinking in young adults: data, definitions, and determinants. Psychol Bull 2009 Jan;135(1):142-156 [FREE Full text] [doi: 10.1037/a0014414] [Medline: 19210057]
6. Boles SM, Miotto K. Substance abuse and violence. Aggression and Violent Behavior 2003 Mar;8(2):155-174. [doi: 10.1016/S1359-1789(01)00057-X]
7. Squeglia LM, Schweinsburg AD, Pulido C, Tapert SF. Adolescent binge drinking linked to abnormal spatial working memory brain activation: differential gender effects. Alcohol Clin Exp Res 2011 Oct;35(10):1831-1841 [FREE Full text] [doi: 10.1111/j.1530-0277.2011.01527.x] [Medline: 21762178]
8. Howland J, Rohsenow DJ, Greece JA, Littlefield CA, Almeida A, Heeren T, et al. The effects of binge drinking on college students' next-day academic test-taking performance and mood state. Addiction 2010 Apr;105(4):655-665 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02880.x] [Medline: 20403018]
9. Mallett KA, Varvil-Weld L, Borsari B, Read JP, Neighbors C, White HR. An update of research examining college student alcohol-related consequences: new perspectives and implications for interventions. Alcohol Clin Exp Res 2013 May;37(5):709-716 [FREE Full text] [doi: 10.1111/acer.12031] [Medline: 23241024]
10. Neighbors C, Atkins DC, Lewis MA, Lee CM, Kaysen D, Mittmann A, et al. Event-specific drinking among college students. Psychol Addict Behav 2011 Dec;25(4):702-707 [FREE Full text] [doi: 10.1037/a0024051] [Medline: 21639597]
11. Dal CS, Worth KA, Gerrard M, Stoolmiller M, Sargent JD, Wills TA, et al. Watching and drinking: expectancies, prototypes, and friends' alcohol use mediate the effect of exposure to alcohol use in movies on adolescent drinking. Health Psychol 2009 Jul;28(4):473-483 [FREE Full text] [doi: 10.1037/a0014777] [Medline: 19594272]
12. Frederiksen NJ, Bakke SL, Dalum P. "No alcohol, no party": an explorative study of young Danish moderate drinkers. Scand J Public Health 2012 Nov;40(7):585-590. [doi: 10.1177/1403494812458988] [Medline: 23027894]
13. Le Garjean N, Vivier-Rousselot V. Conduites addictives: mieux utiliser internet et le mobile en prévention et réduction des risques. 2014. URL: http://www.cirdd-bretagne.fr/fileadmin/publications/Projets/Prevention_internet/docpoursite/V1_Guide_ConduitesAddictives.pdf.Archived [WebCite Cache ID 6P68UrSM8]
14. Facebook. 2012. Newsroom: key facts Internet URL: http://newsroom.fb.com/content/default.aspx?NewsAreaId=22 [WebCite Cache ID 6G2vrxcfo]
15. Capurro D, Cole K, Echavarría MI, Joe J, Neogi T, Turner AM. The use of social networking sites for public health practice and research: a systematic review. J Med Internet Res 2014;16(3):e79 [FREE Full text] [doi: 10.2196/jmir.2679] [Medline: 24642014]
16. Nicholls J. Everyday, everywhere: alcohol marketing and social media--current trends. Alcohol Alcohol 2012 Aug;47(4):486-493 [FREE Full text] [doi: 10.1093/alcalc/ags043] [Medline: 22532575]
17. Ridout B, Campbell A, Ellis L. ‘Off your Face(book)’: alcohol in online social identity construction and its relation to problem drinking in university students. Drug Alcohol Rev 2012 Jan;31(1):20-26. [doi: 10.1111/j.1465-3362.2010.00277.x] [Medline: 21355935]
18. Griffiths R, Casswell S. Intoxigenic digital spaces? Youth, social networking sites and alcohol marketing. Drug Alcohol Rev 2010 Sep;29(5):525-530. [doi: 10.1111/j.1465-3362.2010.00178.x] [Medline: 20887576]
19. Califano JA. Colorado Association of School Boards. 2011. National survey of American attitudes on substance abuse XVI: teens and parents URL: http://www.casb.org/_literature_91867/ERC-NationalStudySubAbuse [accessed 2014-04-25] [WebCite Cache ID 6P682tHVv]
20. Bouthier R. Auvernight: technologies mobiles et internet au service de la prévention [Auvernight: mobile technology and Internet to help health prevention]. 2013 Presented at: Preventions’ Days of INPES; June 5–7, 2013; Paris, France.
21. Goldman M, Darkes J, Del BF. Expectancy mediation of biopsychosocial risk for alcohol use and alcoholism. In: Kirsch EI, editor. How Expectancies Shape Experience. Washington, DC: American Psychological Association; 1999:233-262.
22. Friedman RS, McCarthy DM, Pedersen SL, Hicks JA. Alcohol expectancy priming and drinking behavior: the role of compatibility between prime and expectancy content. Psychol Addict Behav 2009 Jun;23(2):329-333 [FREE Full text] [doi: 10.1037/a0015704] [Medline: 19586149]
23. Stacy AW, Wiers RW. Implicit cognition and addiction: a tool for explaining paradoxical behavior. Annu Rev Clin Psychol 2010;6:551-575 [FREE Full text] [doi: 10.1146/annurev.clinpsy.121208.131444] [Medline: 20192786]
24. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Medical Education 2011 Jun 27;2:53-55. [doi: 10.5116/ijme.4dfb.8dfd]
25. Lin M, Michko G, Bonk C. Characteristics of YouTube use and users: implications for education. 2009 Presented at: World Conference on E-learning in Corporate, Government, Healthcare, and Higher Education; Oct 2009; Vancouver.
26. O’Carroll L. The Guardian. 2012 Dec 18. Twitter active users pass 200 million URL: http://www.guardian.co.uk/technology/2012/dec/18/twitter-users-pass-200-million [accessed 2013-04-21] [WebCite Cache ID 6G2vrxcfv]
27. Orne MT. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist 1962;17(11):776-783. [doi: 10.1037/h0043424]

28. Becoña E, Vázquez FL. Self-reported smoking and measurement of expired air carbon monoxide in a clinical treatment. Psychol Rep 1998 Aug;83(1):316-318. [doi: 10.2466/pr0.1998.83.1.316] [Medline: 9775688]
29. Cherpitel CJ. Alcohol and injuries: a review of international emergency room studies since 1995. Drug Alcohol Rev 2007 Mar;26(2):201-214. [doi: 10.1080/09595230601146686] [Medline: 17364856]
30. Zerhouni O, Bègue L, Brousse G, Carpentier F, Dematteis M, Pennel L, et al. Alcohol and violence in the emergency room: a review and perspectives from psychological and social sciences. Int J Environ Res Public Health 2013 Oct;10(10):4584-4606 [FREE Full text] [doi: 10.3390/ijerph10104584] [Medline: 24084671]
31. Auvernight. Auvernight URL: www.auvernight.org [accessed 2015-12-10] [WebCite Cache ID 6dfdnQn2i]

Abbreviations

SMS: short message service
SNS: social networking site

Edited by G Eysenbach; submitted 14.01.15; peer-reviewed by K Carey, N Johnson; comments to author 02.06.15; revised version received 25.08.15; accepted 15.10.15; published 10.12.15.

Please cite as:
Flaudias V, de Chazeron I, Zerhouni O, Boudesseul J, Begue L, Bouthier R, Lévrier C, Llorca PM, Brousse G
Preventing Alcohol Abuse Through Social Networking Sites: A First Assessment of a Two-Year Ecological Approach
J Med Internet Res 2015;17(12):e278
URL: http://www.jmir.org/2015/12/e278/
doi:10.2196/jmir.4233
PMID:26681577

Original Paper

Is There a Weekly Pattern for Health Searches on Wikipedia and Is the Pattern Unique to Health Topics?

Elia Gabarron[1,2], MSc (Psych); Annie YS Lau[3], Ph.D; Rolf Wynn[2], MD, PhD, Prof

[1]Norwegian Centre for E-health Research, University Hospital of North Norway, Tromsø, Norway

[2]Faculty of Health Sciences, Department of Clinical Medicine, The Arctic University of Norway, Tromsø, Norway

[3]Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

Corresponding Author:
Elia Gabarron, MSc (Psych)
Norwegian Centre for E-health Research
University Hospital of North Norway
Sykehusvegen 23
Tromsø, 9019
Norway
Phone: 47 94863460
Fax: 47 77754099
Email: elia.gabarron@telemed.no

Abstract

Background: Online health information–seeking behaviors have been reported to be more common at the beginning of the workweek. This behavior pattern has been interpreted as a kind of “healthy new start” or “fresh start” due to regrets or attempts to compensate for unhealthy behavior or poor choices made during the weekend. However, the observations regarding the most common health information–seeking day were based only on the analyses of users’ behaviors with websites on health or on online health-related searches. We wanted to confirm if this pattern could be found in searches of Wikipedia on health-related topics and also if this search pattern was unique to health-related topics or if it could represent a more general pattern of online information searching—which could be of relevance even beyond the health sector.

Objective: The aim was to examine the degree to which the search pattern described previously was specific to health-related information seeking or whether similar patterns could be found in other types of information-seeking behavior.

Methods: We extracted the number of searches performed on Wikipedia in the Norwegian language for 911 days for the most common sexually transmitted diseases (chlamydia, gonorrhea, herpes, human immunodeficiency virus [HIV], and acquired immune deficiency syndrome [AIDS]), other health-related topics (influenza, diabetes, and menopause), and 2 nonhealth-related topics (footballer Lionel Messi and pop singer Justin Bieber). The search dates were classified according to the day of the week and ANOVA tests were used to compare the average number of hits per day of the week.

Results: The ANOVA tests showed that the sexually transmitted disease queries had their highest peaks on Tuesdays ($P<.001$) and the fewest searches on Saturdays. The other health topics also showed a weekly pattern, with the highest peaks early in the week and lower numbers on Saturdays ($P<.001$). Footballer Lionel Messi had the highest mean number of hits on Tuesdays and Wednesdays, whereas pop singer Justin Bieber had the most hits on Tuesdays. Both these tracked search queries also showed significantly lower numbers on Saturdays ($P<.001$).

Conclusions: Our study supports prior studies finding an increase in health information searching at the beginning of the workweek. However, we also found a similar pattern for 2 randomly chosen nonhealth-related terms, which may suggest that the search pattern is not unique to health-related searches. The results are potentially relevant beyond the field of health and our preliminary findings need to be further explored in future studies involving a broader range of nonhealth-related searches.

(J Med Internet Res 2015;17(12):e286) doi:10.2196/jmir.5038

KEYWORDS

information-seeking behavior; health information–seeking behavior; periodicity; Wikipedia; chlamydia; gonorrhea; HIV; AIDS; influenza; diabetes

Introduction

People tend to structure their activities in a weekly (circaseptan) pattern [1], with some days for work and some for rest. The social construct of a 7-day cycle is not new and seems to have its origin around 500 BC [1]. In modern society, many behaviors typically follow a weekly pattern, as do many somatic and psychological symptoms and disorders. For instance, heart attacks, strokes, and migraines tend to be most frequent during the workweek [2-4]. These patterns may be related to differences in physical activity, blood pressure, and stress levels between the workweek and days off from work. Other health activities and aspirational behaviors, such as attending gym, being on time, or quitting smoking, have also been found to be more frequent at the beginning of the week and after temporal landmarks (ie, relocation, job change, birthday, first day of spring). This has been called the “fresh start effect” [5,6].

Sleeping is another central behavior that follows a weekly pattern, typically with a lack of sleep during the workweek (“social jet lag”) and compensatory sleeping on days off [7,8]. Mood and level of aggression also follow various cycles; for instance, suicides are most frequent on Mondays and panic attacks are most frequent on days off from work [9-12].

Sexual risk behaviors, which typically are related to drugs or alcohol consumption, have been reported to be more frequent during the weekends [13,14], whereas online health information-seeking behaviors seem to be more common at the beginning of the workweek [14-16]. Interestingly, the increased rates in online health information-seeking behavior at the beginning of the week have been interpreted as regrets or attempts to compensate for unhealthy behavior or poor choices made during the weekend [14-16] and also as a kind of “healthy new start” [17,18] in agreement with the “fresh start effect” hypotheses [6]. These hypotheses explain the increased information-seeking activity at the beginning of the week as an “aspirational behavior.” This implies that these days (ie, at the beginning of the workweek) would be when people are most motivated to pursue their aspirations or most likely to think about their health [6,17,18].

In this sense, a Spanish website on sexual health found peaks in the number of consultations received by phone and email on Mondays and Tuesdays [15] and a Dutch website on sexual health reported peaks in their visitor rates every Monday [16]. The searches performed on Google seem to follow the same weekly pattern; analyses of the number of US searches including the term “healthy” or “diet” as well as smoking cessation queries showed peaks at the beginning of the week [6,17,18]. The analyses of the number of postings on online smoking cessation support networks also have weekly patterns, with highest activity during the workweek and lower numbers of postings on Saturdays [19].

In these studies, the observations regarding the most common health information-seeking day are based only on the analyses of users’ behaviors with websites on health or on health-related searches on the Google engine. We wanted to examine if the pattern described previously could also be observed in searches on Wikipedia and to what degree the search pattern was specific to health-related information seeking or whether similar patterns could be found in nonhealth-related information-seeking behavior. Therefore, to test this idea, we hypothesized that the information-seeking pattern with peaks in searches at the beginning of the workweek was specific to health-related information seeking. If a similar pattern for nonhealth-related topics was also found, this could mean that a more general pattern of online information searching existed—which could be of relevance even beyond the health sector. The information-seeking behavior shown by people who search Wikipedia can be considered to be representative for online information-seeking behavior patterns in general. Qualitative studies have found that accessibility, perceived trustworthiness, and usability are the most important criteria for online health information seekers [20,21]. Moreover, Wikipedia might represent one of the most frequently used online resources for information seeking and health information seeking in countries with high Internet penetration rates, such as Norway [22]. Wikipedia appears on the first page of Google search results, it is considered a trustable source of information on health [23,24], and the information on Wikipedia is even used by health professionals and researchers [25].

Methods

Wikipedia has become one of the main sources of information on the Internet [26,27]. Its excellent Web positioning / search engine optimization makes Wikipedia one of the first hits after searching almost any word on the Internet. Currently, it has more than 35 million articles written in 288 different languages and the Norwegian edition, with 413,459 articles, represents the 19th largest language edition [27]. By the end of September 2015, the Norwegian Wikipedia edition had 58,706 page views per hour. Hypothetically, if each of these searches were performed by different individuals, this could imply that 28% of Norwegians accessed Wikipedia daily [28]. Although this figure is likely too high, the point stands that Wikipedia has become a central source of online information for Norwegians.

To examine the search patterns related to health-related topics and nonhealth-related topics, we tracked the number of searches performed in the Norwegian language on Wikipedia from January 1, 2013 to June 30, 2015.

The traffic statistics was extracted from 10 Wikipedia articles. Because there have been several prior publications on the weekly pattern of information seeking about sexually transmitted diseases (STDs), we chose to examine searches for 5 of the most common STDs (chlamydia, gonorrhea, herpes, human immunodeficiency virus [HIV], and acquired immune deficiency syndrome [AIDS]). To get a broader picture of the search pattern for diseases, we also chose to include a common seasonal infectious disease (influenza), which could present a more seasonal pattern. We also chose to include a common noninfectious disease (diabetes) [29] as well as a natural bodily process (menopause) to further broaden the scope of the searches. To test our hypothesis that the search pattern with peaks at the beginning of the workweek would be specific to health-related information searching, we included 2 randomly chosen celebrities in sports and music: 3-time FIFA Ballon d’Or winner Lionel Messi and one of the current top-selling pop stars, Justin Bieber. Thus, the articles we tracked in this study were chlamydia, gonorrhea, herpes, HIV, AIDS, influenza, diabetes, menopause, Justin Bieber, and Lionel Messi.

The daily rates of Wikipedia article hits were extracted from the Wikipedia article traffic statistics website [30]. This website counts Wikipedia page views per day and classifies the views by the article titles [30]. The dates were classified according to the day of the week. Public holidays in Norway and days after public holidays were also identified.

Descriptive statistics were used to summarize the absolute numbers and frequencies of hits per day. ANOVA tests were used to compare the means of hits per day of the week. The Mann-Whitney *U* test was used to compare the means of hits during public holidays and the first day after the public holidays. Data were analyzed with SPSS version 22.

Results

A total of 10 articles on Wikipedia were tracked for 911 days (from January 1, 2013 to June 30, 2015). All the tracked Wikipedia hits showed a significant weekly pattern with highest peaks early in the week, mostly on Mondays and Tuesdays, and lower numbers on Saturdays. The daily mean queries per search term and their 95% confidence intervals are summarized in Table 1 and Figures 1-3.

Table 1. Mean weekly Wikipedia traffic (January 1, 2013-June 30, 2015).

Wikipedia information searches	Day of week, mean (95% CI)							*P* [a]
	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday	
Sexual diseases								
Chlamydia	76.8 (65.2-88.5)	81.5 (69.0-94.0)	74.5 (63.4-85.6)	73.8 (62.8-84.7)	58.9 (50.3-67.5)	37.1 (31.4-42.7)	50.1 (42.3-57.9)	<.001
Gonorrhea	24.4 (21.8-27.0)	28.5 (24.5-32.4)	28.0 (24.5-31.5)	25.9 (23.4-28.5)	21.5 (19.5-23.5)	16.5 (14.8-18.3)	18.4 (16.5-20.2)	<.001
Herpes	60.4 (49.7-71.1)	67.6 (44.7-90.5)	55.0 (48.4-61.7)	51.7 (47.4-56.1)	46.0 (42.5-49.5)	37.3 (34.1-40.5)	44.5 (40.8-48.2)	<.001
HIV	101.9 (92.4-111.3)	106.1 (94.7-117.4)	98.7 (90.5-106.9)	90.9 (83.9-97.9)	77.5 (72.0-83.1)	55.0 (50.7-59.3)	66.9 (62.0-71.9)	<.001
AIDS	79.5 (69.6-89.5)	85.5 (74.9-96.2)	75.4 (69.3-81.5)	73.8 (67.8-79.8)	62.8 (58.3-67.2)	45.4 (36.5-54.3)	53.2 (48.3-58.2)	<.001
Other health topics								
Influenza	58.2 (47.5-68.9)	58.2 (46.2-70.1)	59.8 (43.9-75.7)	56.9 (41.9-71.8)	45.7 (34.0-57.5)	34.7 (27.0-42.3)	44.0 (34.4-53.7)	<.001
Diabetes	130.4 (117.2-143.5)	136.8 (123.4-150.1)	136.5 (123.7-149.2)	128.9 (116.2-141.5)	98.4 (89.9-106.9)	58.6 (54.2-63.0)	83.6 (76.0-91.2)	<.001
Menopause	48.4 (45.2-51.7)	47.7 (44.2-51.2)	47.3 (43.8-50.9)	45.1 (41.6-48.7)	39.0 (36.4-41.6)	33.9 (31.3-36.6)	42.9 (39.7-46.2)	<.001
Nonhealth-related topics								
Lionel Messi	203.1 (179.3-226.5)	227.7 (204.5-250.9)	228.1 (203.4-252.9)	199.1 (181.8-216.3)	148.8 (135.6-161.9)	124.1 (108.4-139.9)	151.0 (117.7-184.3)	<.001
Justin Bieber	371.9 (279.3-464.6)	408.7 (304.5-512.9)	395.9 (303.1-488.7)	389.8 (299.0-480.6)	337.8 (253.2-422.3)	244.0 (176.8-311.2)	258.1 (192.4-323.8)	.03

[a] ANOVA.

The ANOVA tests showed that the Wikipedia queries on STDs performed in Norwegian had their highest peaks on Tuesdays ($P<.001$) and Saturday was the day with the fewest searches on STDs. Of the STD searches examined, HIV and AIDS were most frequently searched for. For instance, on Tuesdays, there were a mean 106.1 (95% CI 94.7-117.4) hits on HIV and 85.5 (95% CI 74.9-96.2) hits on AIDS. Comparing the number of hits on Tuesdays and Saturdays, the number of hits was 54.5% lower for searches on chlamydia and 48.2%, 46.9%, 44.8%, and 42.1% lower for HIV, AIDS, herpes, and gonorrhea, respectively.

The other health topics searched on Wikipedia also showed a weekly pattern, with the highest peaks early in the week and lower numbers on Saturdays ($P<.001$). The influenza queries had their highest frequency from Monday to Wednesday, with a peak of 59.8 hits on Wednesdays, on average. The diabetes queries were most frequent on Tuesdays and Wednesdays (peaks of 136.8 hits and 136.5 hits, respectively). The term menopause was most frequently searched on Mondays and Tuesdays (mean peaks of 48.4, 95% CI 45.2-51.7 hits and 47.7, 95% CI 44.2-51.2 hits, respectively). The 3 other health topics tracked showed significantly lower numbers on Saturdays ($P<.001$). The number of queries on Wikipedia decreased 29.9% for searches on menopause, 41.9% for influenza, and 57.2% for diabetes between the days with the highest and lowest numbers of hits. See Table 1 and Figure 2.

Barcelona footballer Lionel Messi had the highest mean number of hits on Tuesdays and Wednesdays (227.7, 95% CI

204.5-250.9 and 228.1, 95% CI 203.4-252.9, respectively), whereas celebrity pop singer Justin Bieber had the most hits on Tuesdays (mean 408.7, 95% CI 304.5-512.9). Both tracked search queries also showed significantly lower numbers on Saturdays, with a decrease of 45.6% for Messi and 40.3% for Bieber. See Table 1 and Figure 3.

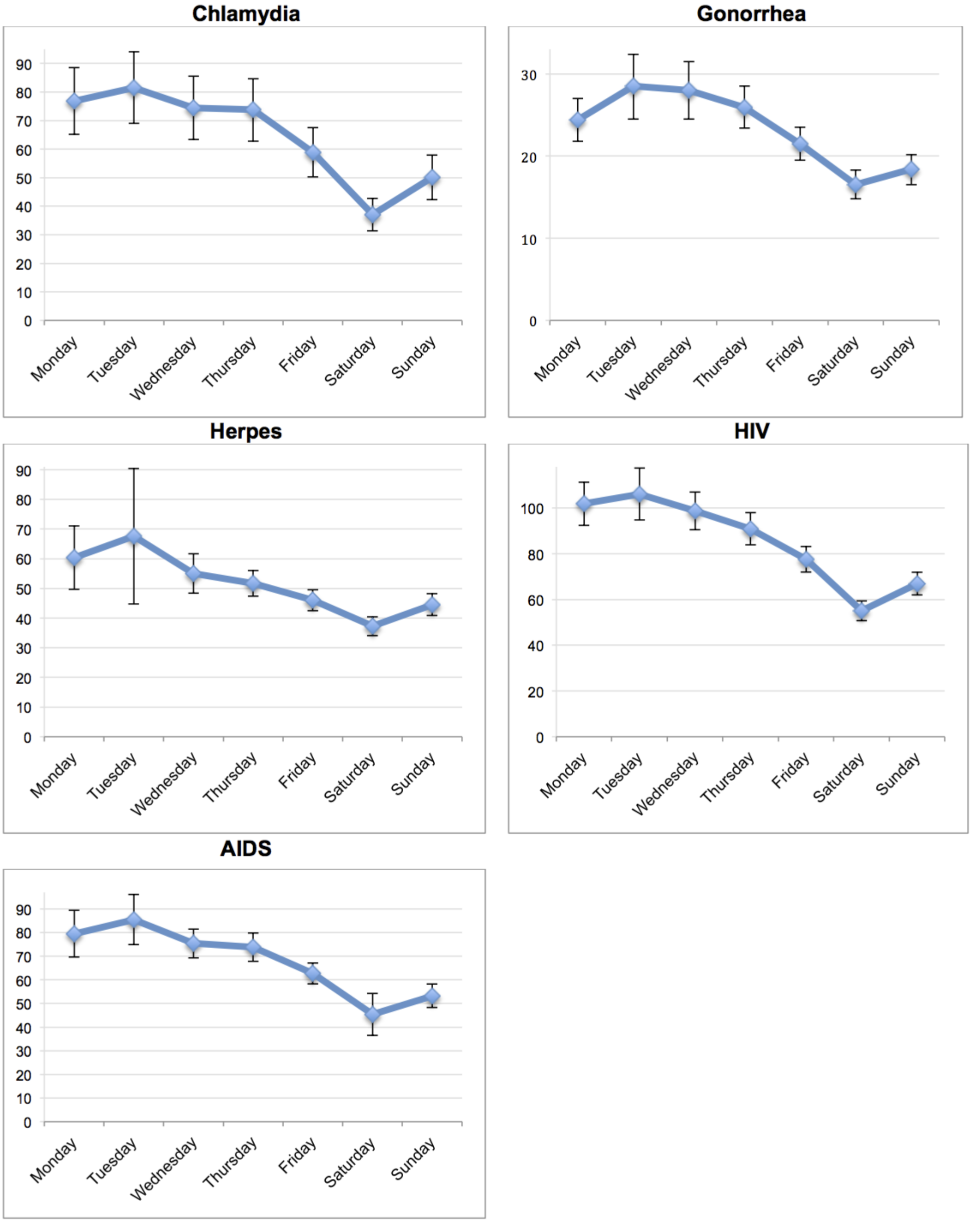
A total of 37 public holidays in Norway and 18 days following the public holidays were identified. Mann-Whitney *U* tests showed significant mean differences for online searches performed on public holidays and days after holidays only for the searches on HIV (P=.01), AIDS (P=.03), diabetes (P=.01), and Lionel Messi (P=.04). No differences were found in the remaining online searches analyzed. The mean and 95% CI of hits on Wikipedia during public holidays and the days after public holidays are summarized in Table 2.

Table 2. Wikipedia traffic on public holidays and days after public holidays.

Wikipedia information searches	Public holidays, mean (95% CI) n=37	Days after public holidays, mean (95% CI) n=18	P [a]
Sexual diseases			
Chlamydia	43.3 (29.3-57.3)	66.5 (31.9-101.1)	.30
Gonorrhea	17.7 (12.7-22.7)	22.1 (16.0-28.1)	>.99
Herpes	40.8 (33.5-48.1)	55.5 (37.3-73.7)	.11
HIV	56.0 (47.8-64.2)	81.1 (63.0-99.2)	.01
AIDS	41.4 (35.3-47.6)	61.2 (44.3-78.1)	.03
Other health topics			
Influenza	41.4 (15.4-67.3)	81.9 (-17.6-184.5)	.30
Diabetes	69.1 (58.8-79.3)	114.9 (78.0-151.9)	.01
Menopause	38.2 (31.8-44.7)	44.5 (34.6-54.4)	.24
Nonhealth-related topics			
Lionel Messi	111.2 (85.1-137.3)	170.0 (118.6-221.4)	.04
Justin Bieber	229.2 (123.1-335.3)	327.8 (134.3.1-521.4)	.47

[a] Mann-Whitney *U* test (2-tailed).

Figure 1. Weekly mean online information searches on sexual diseases (January 1, 2013-June 30, 2015). Error bars indicate 95% CI. All ANOVA tests P<.001.



Chlamydia

Gonorrhea

Herpes

HIV

AIDS

Figure 2. Weekly mean online information searches on other health topics (January 1, 2013-June 30, 2015). Error bars indicate 95% CI. All ANOVA tests P<.001.

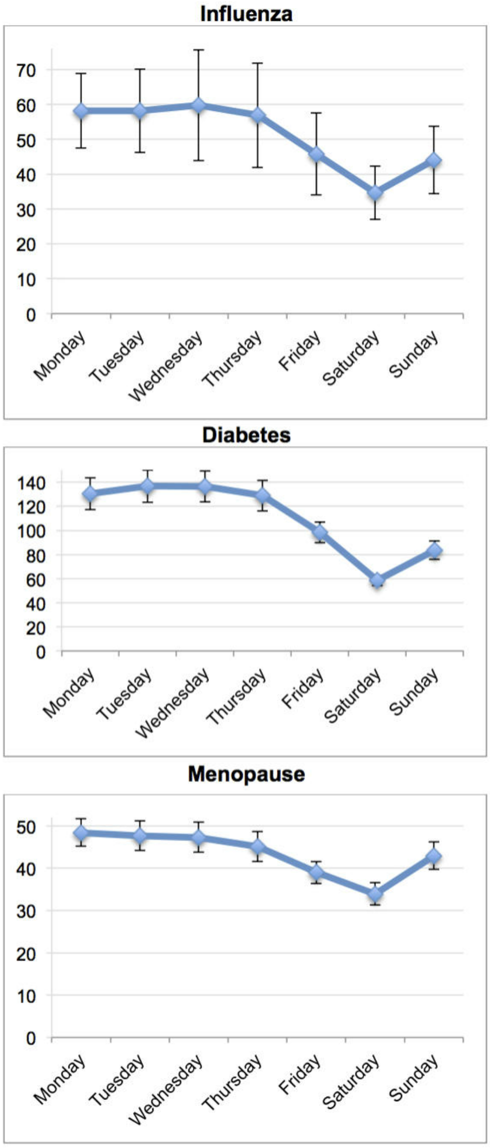
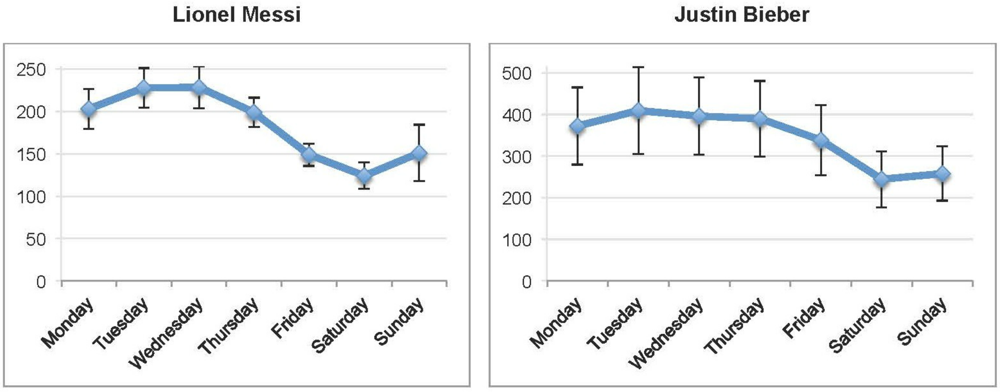


Figure 3. Weekly mean online information searches on nonhealth-related topics (January 1, 2013-June 30, 2015). Error bars indicate 95% CI. All ANOVA tests P<.05.



Discussion

Our results show that all online queries we examined followed a circaseptan, or weekly, pattern independent of the nature of the search query (sexual diseases, other health topics, or nonhealth-related topics). This weekly online information-seeking pattern has its higher peaks early in the week, mostly on Mondays and Tuesdays, and its low peaks on

Saturdays. To our knowledge, this is the first study suggesting the possibility of a weekly pattern in general online information-seeking behavior relating to topics beyond health.

A Weekly Online Information-Seeking Pattern

Online information-seeking behavior is an intentional, planned behavior that is performed on the Internet. In the case of online searches related to health, one could speculate that the searches are performed mostly by adults with a specific interest in the topics, maybe because they or a loved one have been diagnosed with a condition or because they think they might have a health-related problem [31-33]. Some previous studies have reported weekly patterns of health information-seeking behavior; these findings have been interpreted by some as reflecting people regretting their unhealthy weekend behaviors [14-16] and also as a kind of "healthy new start" or "fresh start" [6,17,18]. Interestingly, in our results, the nonhealth-related queries also followed the same weekly pattern and this implies that the pattern is not unique to health-related searching and there is a need to further examine if online early week information-seeking behavior could represent a more general information-seeking behavior pattern. Furthermore, it is possible that these information-seeking patterns could apply to various age groups because disease-related searches were more likely to have been performed by adults, whereas most searches on the pop singer Justin Bieber were probably performed by teenagers (ie, his main fan base).

Possibly, the online information-seeking behavior pattern could be understood in light of the "fresh start effect" hypothesis in the sense that the higher number of searches on Wikipedia could result from an increased motivation of people to increase their knowledge by seeking information at the beginning of the week [5,6]. However, in our sample, the most common days to perform the online searches were Mondays and Tuesdays, whereas the "fresh start effect" hypothesis would suggest the highest frequency on Mondays only. If we consider the public holidays as a temporal landmark, our results do not fit the "fresh start effect" hypothesis completely. Although the mean number of hits performed on Wikipedia on days after public holidays is higher than the mean number of hits on public holidays, our results show, in all cases, these numbers are lower than the ones observed at the beginning of the week. This could mean that the start of the week has a more powerful effect than temporal landmarks, such as public holidays. On the other hand, the mean number of online hits on Saturdays was, in all our observations, lower than the number of searches performed on public holidays.

Do New Technologies Have an Effect on the Weekly Behavior Pattern?

In a world where people have constant access to the Internet, it could be that we tend to structure our online information-seeking activities in a weekly pattern; information-seeking behavior activities occur more frequently early in the week, whereas the least popular days for online information searching are Saturdays. In the same way that there are days for work and days for rest, there might be days when online information seeking is more common and days when seeking information online is less frequent. We have not found publications analyzing the existence of information-seeking behavior patterns in the offline world; therefore, we cannot say if these online patterns are just a reflection of usual offline information-seeking behavior. Maybe the weekly online information-seeking pattern might respond to new behavior routines related to the appearance of the new technologies. And, if so, some other behaviors linked to the use of the new technologies could also follow a temporal pattern.

Relevance of Online Information for Decision Making, Policy Making, and for Public and Private Institutions

Much of modern life is organized in weekly patterns and these patterns are of importance to behavior and to when many symptoms and illnesses occur. Although we are unsure about the underlying causes driving the weekly online information-seeking pattern, we believe the pattern might be relevant to those who wish to reach a largest possible audience with specific information, for instance. If the pattern we described is generic for other sources of online information, it might be important to consider which day of the week that is chosen to release important information to get maximum online exposure. Thus, if confirmed in future studies, the search pattern information could be relevant for press releases or public campaigns of public health agencies, nongovernmental organizations, charities, online publishers, etc. When dealing with this issue, it should also be taken into consideration that people may also be more likely to search for other types of information (ie, not related to health) on the same days. This implies that it will remain a challenge to make the health-related message or campaign attract attention and stand out from the other information that is available.

Increasingly, online information plays an important part in decision making [34]. Decisions are improved by better access to relevant information and searching for documents on the Web is increasingly an important source of that information [35]. However, although past research has focused on population access and usage of the Internet [34,36,37], there has been little, if any, examination of when general information needs or motivations (in addition to health information-seeking behavior) are likely to arise during the week, whether searches are conducted in temporal patterns, and how they influence the actual decisions that are being made in an everyday context [38]. This information could be important in the development of integrated information retrieval systems that support decision making. In agreement with the "fresh start effect" studies, we believe that the implications of these findings may be important for policy making and for public and private institutions targeting the general public because campaigns or messages could have the greatest impact at the beginning of the week when people seem to be most eagerly searching for information [6]. However, more research, including empirical testing, is needed before stronger conclusions can be made.

Limitations

Our study tracked information regarding 10 article queries performed on Wikipedia in the Norwegian language. This means our findings should be regarded as preliminary and further research should be undertaken to check if these weekly patterns can be found for online searching behavior in general. This implies examining if the pattern stands for a broader range of search terms and in different languages and cultures. Moreover, because Wikipedia searches are anonymous, we do not know how variables such as age, gender, health status, place of living, employment status, etc, impact the searching pattern. These and other variables may be particularly important when conducting public health campaigns. It is also of interest to see to what extent a similar pattern of information seeking can be found for other sources of information, such as social media. It will be interesting to explore the existence of other temporal behavior patterns in relation to the appearance of new technologies and their potential impact on decision making. Although it would be interesting to test our findings in an actual health campaign, this lies beyond the scope of this study.

Conclusion

Our study supported prior studies finding an increase in health information searching at the beginning of the workweek. However, we also found a similar pattern for 2 randomly chosen nonhealth-related terms, which may suggest that the search pattern is not unique to health-related searches. The results are potentially relevant beyond the field of health and our preliminary findings need to be further explored in future studies involving a broader range of nonhealth-related searches.

Acknowledgments

This work was supported by the North Norway Regional Health Authority (Helse Nord RHF), grant number HST1184-14.

Authors' Contributions

EG had the idea for the paper, gathered and analyzed the data, drafted the paper, revised the paper, and approved the final version. AYSL contributed to the Discussion, reviewed/edited the manuscript, and approved the final version. RW had the idea for the paper, reviewed/edited the paper, and approved the final version.

Conflicts of Interest

None declared.

References

1. Zerubavel E. The Seven Day Circle: The History and Meaning of the Week. Chicago: University of Chicago Press; 1985.
2. Manfredini R, Fabbian F, Pala M, Tiseo R, De Giorgi A, Manfredini F, et al. Seasonal and weekly patterns of occurrence of acute cardiovascular diseases: does a gender difference exist? J Womens Health (Larchmt) 2011 Nov;20(11):1663-1668. [doi: 10.1089/jwh.2011.2734] [Medline: 21967079]
3. Turin T, Kita Y, Rumana N, Takashima N, Ichikawa M, Sugihara H, et al. Circaseptan variation in case-fatality rate for patients with acute subarachnoid hemorrhage (Takashima Stroke Registry 1988-2003). J Clin Neurosci 2010 Jul;17(7):869-873. [doi: 10.1016/j.jocn.2009.11.004] [Medline: 20395145]
4. Alstadhaug K, Salvesen R, Bekkelund S. Weekend migraine. Cephalalgia 2007 Apr;27(4):343-346. [doi: 10.1111/j.1468-2982.2007.01284.x] [Medline: 17376111]
5. Dai H, Milkman K, Riis J. Put your imperfections behind you: why and how meaningful temporal landmarks motivate aspirational behavior. SSRN Journal 2014 Jun 09:e1 [FREE Full text] [doi: 10.2139/ssrn.2420476]
6. Dai H, Milkman K, Riis J. The fresh start effect: temporal landmarks motivate aspirational behavior. Manage Sci 2014 Oct;60(10):2563-2582. [doi: 10.1287/mnsc.2014.1901]
7. Johnsen M, Wynn R, Allebrandt K, Bratlid T. Lack of major seasonal variations in self reported sleep-wake rhythms and chronotypes among middle aged and older people at 69 degrees North: the Tromsø Study. Sleep Med 2013 Feb;14(2):140-148. [doi: 10.1016/j.sleep.2012.10.014] [Medline: 23219143]
8. Friborg O, Rosenvinge JH, Wynn R, Gradisar M. Sleep timing, chronotype, mood, and behavior at an Arctic latitude (69° N). Sleep Med 2014 Jul;15(7):798-807. [doi: 10.1016/j.sleep.2014.03.014] [Medline: 24933084]
9. Kao L, Xirasagar S, Chung K, Lin H, Liu S, Chung S. Weekly and holiday-related patterns of panic attacks in panic disorder: a population-based study. PLoS One 2014;9(7):e100913 [FREE Full text] [doi: 10.1371/journal.pone.0100913] [Medline: 25006664]
10. McCleary R, Chew KS, Hellsten JJ, Flynn-Bransford M. Age- and sex-specific cycles in United States suicides, 1973 to 1985. Am J Public Health 1991 Nov;81(11):1494-1497. [Medline: 1951813]
11. Wynn R. Polar day and polar night: month of year and time of day and the use of physical and pharmacological restraint in a north Norwegian university psychiatric hospital. Arctic Med Res 1996 Oct;55(4):174-181. [Medline: 9115543]
12. Cornélissen G, Watson D, Mitsutake G, Fišer B, Siegelová J, Dušek J, Vohlídalová, et al. Mapping of circaseptan and circadian changes in mood. Scr Med (Brno) 2005;78(2):89-98 [FREE Full text] [Medline: 18985163]
13. Fortenberry J, Orr D, Zimet G, Blythe M. Weekly and seasonal variation in sexual behaviors among adolescent women with sexually transmitted diseases. J Adolesc Health 1997 Jun;20(6):420-425. [doi: 10.1016/S1054-139X(96)00275-3] [Medline: 9178078]

14. Fuentes L, Johnson M, Holtgrave D. An exploration of weekly patterns in HIV-related behaviors: implications for successful interventions and future research. AIDS Care 2015 Sep;27(9):1118-1127. [doi: 10.1080/09540121.2015.1032204] [Medline: 25894315]
15. Marco J, Barba R, Losa JE, de la Serna Carlos Martínez, Sainz M, Lantigua IF, de la Serna Jose Luis. Advice from a medical expert through the Internet on queries about AIDS and hepatitis: analysis of a pilot experiment. PLoS Med 2006 Jul;3(7):e256 [FREE Full text] [doi: 10.1371/journal.pmed.0030256] [Medline: 16796404]
16. Crutzen R, Mevissen FE. Just another manic Monday: peaking sexual concerns after the weekend. Arch Sex Behav 2011 Dec;40(6):1105-1106 [FREE Full text] [doi: 10.1007/s10508-011-9833-2] [Medline: 21837516]
17. Ayers J, Althouse B, Johnson M, Cohen J. Circaseptan (weekly) rhythms in smoking cessation considerations. JAMA Intern Med 2014 Jan;174(1):146-148 [FREE Full text] [doi: 10.1001/jamainternmed.2013.11933] [Medline: 24166181]
18. Ayers J, Althouse B, Johnson M, Dredze M, Cohen J. What's the healthiest day?: Circaseptan (weekly) rhythms in healthy considerations. Am J Prev Med 2014 Jul;47(1):73-76. [doi: 10.1016/j.amepre.2014.02.003] [Medline: 24746375]
19. Healey B, Hoek J, Edwards R. Posting behaviour patterns in an online smoking cessation social network: implications for intervention design and development. PLoS One 2014;9(9):e106603 [FREE Full text] [doi: 10.1371/journal.pone.0106603] [Medline: 25192174]
20. Zhang Y. Beyond quality and accessibility: Source selection in consumer health information searching. J Assoc Inf Sci Technol 2014;65(5):911-927.
21. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
22. Statistics Norway. 2015. ICT usage in households, 2014, 2nd quarter URL: http://www.ssb.no/en/ikthus [accessed 2015-08-14] [WebCite Cache ID 6am1qSlRG]
23. Maskalyk J. Modern medicine comes online: how putting Wikipedia articles through a medical journal's traditional process can put free, reliable information into as many hands as possible. Open Med 2014;8(4):e116-e119 [FREE Full text] [Medline: 25426179]
24. Kräenbring J, Monzon PT, Gutmann J, Muehlich S, Zolk O, Wojnowski L, et al. Accuracy and completeness of drug information in Wikipedia: a comparison with standard textbooks of pharmacology. PLoS One 2014;9(9):e106930 [FREE Full text] [doi: 10.1371/journal.pone.0106930] [Medline: 25250889]
25. Bould M, Hladkowicz E, Pigford A, Ufholz L, Postonogova T, Shin E, et al. References that anyone can edit: review of Wikipedia citations in peer reviewed health science literature. BMJ 2014;348:g1585 [FREE Full text] [Medline: 24603564]
26. Heilman J, Kemmann E, Bonert M, Chatterjee A, Ragar B, Beards G, et al. Wikipedia: a key tool for global public health promotion. J Med Internet Res 2011;13(1):e14 [FREE Full text] [doi: 10.2196/jmir.1589] [Medline: 21282098]
27. Wikipedia. 2015. Wikipedia:About URL: https://en.wikipedia.org/wiki/Wikipedia:About [accessed 2015-08-14] [WebCite Cache ID 6am1mntu5]
28. Wikimedia. 2015. Wikipedia Statistics - 30 September URL: http://stats.wikimedia.org/EN/Sitemap.htm [accessed 2015-10-20] [WebCite Cache ID 6cPpK6lH9]
29. World Health Organization Regional Office for Europe. 2015. Diabetes URL: http://www.euro.who.int/en/health-topics/noncommunicable-diseases/diabetes/diabetes [accessed 2015-10-22] [WebCite Cache ID 6cSlgX29j]
30. Wikipedia. 2015. Wikipedia article traffic statistics URL: http://stats.grok.se/ [accessed 2015-10-14] [WebCite Cache ID 6cGizsdWA]
31. Bratucu R, Gheorghe I, Purcarea R, Gheorghe C, Popa VO, Purcarea V. Cause and effect: the linkage between the health information seeking behavior and the online environment--a review. J Med Life 2014 Sep 15;7(3):310-316 [FREE Full text] [Medline: 25408746]
32. Chisolm D. Does online health information seeking act like a health behavior?: A test of the behavioral model. Telemed J E Health 2010 Mar;16(2):154-160. [doi: 10.1089/tmj.2009.0102] [Medline: 20156127]
33. Santana S, Lausen B, Bujnowska-Fedak M, Chronaki C, Prokosch H, Wynn R. Informed citizen and empowered citizen in health: results from an European survey. BMC Fam Pract 2011;12:20 [FREE Full text] [doi: 10.1186/1471-2296-12-20] [Medline: 21496309]
34. Americans Feel Better Informed Thanks to the Internet. Washington, DC: Pew Internet & American Life Project; 2014 Dec 08. URL: http://www.pewinternet.org/files/2014/12/PI_InformedWeb_120814_02.pdf [accessed 2015-08-14] [WebCite Cache ID 6am1uarhk]
35. Morrison J, Pirolli P, Card S. A taxonomic analysis of what world wide web activities significantly impact people’s decisions and actions. In: Proceedings of CHI EA '01 CHI '01 Extended Abstracts on Human Factors in Computing Systems. 2001 Presented at: Conference on Human Factors in Computing System; Mar 31-Apr 05, 2001; Seattle, WA. [doi: 10.1145/634067.634167]
36. Perrin A, Duggan M. Americans' Internet Access: 2000-2015. Washington, DC: Pew Internet & American Life Project; 2015 Jun 26. URL: http://www.pewinternet.org/files/2015/06/2015-06-26_internet-usage-across-demographics-discover_FINAL.pdf [accessed 2015-08-14] [WebCite Cache ID 6am1xeVN8]

37. Kummervold PE, Wynn R. Health information accessed on the internet: the development in 5 European countries. Int J Telemed Appl 2012;2012:297416 [FREE Full text] [doi: 10.1155/2012/297416] [Medline: 23304133]
38. Hersh WR. Ubiquitous but unfinished: online information retrieval systems. Med Decis Making 2005;25(2):147-148. [doi: 10.1177/0272989X05275557] [Medline: 15800298]

Abbreviations

AIDS: acquired immune deficiency syndrome
HIV: human immunodeficiency virus
STD: sexually transmitted disease

Edited by G Eysenbach; submitted 26.08.15; peer-reviewed by V Dedov, Y Albalawi, M Bujnowska-Fedak; comments to author 09.10.15; revised version received 28.10.15; accepted 29.11.15; published 22.12.15.

Please cite as:
Gabarron E, Lau AYS, Wynn R
Is There a Weekly Pattern for Health Searches on Wikipedia and Is the Pattern Unique to Health Topics?
J Med Internet Res 2015;17(12):e286
URL: http://www.jmir.org/2015/12/e286/
doi:10.2196/jmir.5038
PMID:26693859

Original Paper

Evaluation of a Web-Based Food Record for Children Using Direct Unobtrusive Lunch Observations: A Validation Study

Anine Christine Medin[1], RD, MS; Helene Astrup[1], MS; Britt Marlene Kåsin[2], MS; Lene Frost Andersen[1], PhD

[1]Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Oslo, Norway

[2]Institute of Health and Society, Faculty of Medicine, University of Oslo, Oslo, Norway

Corresponding Author:
Anine Christine Medin, RD, MS
Institute of Basic Medical Sciences
Faculty of Medicine
University of Oslo
Postboks 1046
Blindern
Oslo, 0317
Norway
Phone: 47 22851349
Fax: 47 22851249
Email: a.c.medin@medisin.uio.no

Abstract

Background: High-quality, Web-based dietary assessment tools for children are needed to reduce cost and improve user-friendliness when studying children's dietary practices.

Objective: To evaluate the first Web-based dietary assessment tool for children in Norway, the Web-based Food Record (WebFR), by comparing children's true school lunch intake with recordings in the WebFR, using direct unobtrusive observation as the reference method.

Methods: A total of 117 children, 8-9 years, from Bærum, Norway, were recruited from September to December 2013. Children completed 4 days of recordings in the WebFR, with parental assistance, and were observed during school lunch in the same period by 3 observers. Interobserver reliability assessments were satisfactory. Match, omission, and intrusion rates were calculated to assess the quality of the recordings in the WebFR for different food categories, and for all foods combined. Logistic regression analyses were used to investigate whether body mass index (BMI), parental educational level, parental ethnicity or family structure were associated with having a "low match rate" (≤70%).

Results: Bread and milk were recorded with less bias than spreads, fruits, and vegetables. Mean (SD) for match, omission, and intrusion rates for all foods combined were 73% (27%), 27% (27%), and 19% (26%), respectively. Match rates were statistically significantly associated with parental educational level (low education 52% [32%] versus high 77% [24%], *P*=.008) and parental ethnicity (non-Norwegian 57% [28%] versus others 75% [26%], *P*=.04). Only parental ethnicity remained statistically significant in the logistic regression model, showing an adjusted odds ratio of 6.9 and a 95% confidence interval between 1.3 and 36.4.

Conclusions: Compared with other similar studies, our results indicate that the WebFR is in line with, or better than most of other similar tools, yet enhancements could further improve the WebFR.

(J Med Internet Res 2015;17(12):e273) doi:10.2196/jmir.5031

KEYWORDS

children; dietary records; Internet; observation; validity

Introduction

High-quality dietary assessment tools are essential when studying children's dietary practices. Traditional tools, such as food frequency questionnaires, 24-hour recalls, and food records, can be used to assess dietary intake [1-3]. In recent years, there has been a shift toward the use of Web-based dietary assessment tools among both adults and the younger age groups [4-6]; those aimed at children are mostly 24-hour recalls, or mixed methods combining food records and 24-hour recalls [7-12]. These new types of digital dietary assessment methods are highly needed [3].

In comparison with paper-based dietary assessment tools, Web-based tools facilitate data handling and improve user-friendliness; they reduce the burden for both the participant and researcher and can enhance motivation [3,6]. Therefore, we have recently adapted the Danish Web-based Dietary Assessment Software for Children (WebDASC) [7] to Norwegian conditions and food culture to develop the Web-based Food Record (WebFR) for children and adolescents.

It is well established that assessment of dietary intake is associated with errors [13]. Furthermore, assessing children's intake is especially challenging due to their limited cognitive abilities [14], and because they often need assistance from a caretaker [13]. Consequently, validating dietary assessment tools that target children is important [1].

Direct observation is considered to be an appropriate high-quality method for validation studies of dietary assessment tools, because it provides direct unbiased information regarding what is eaten [15,16]. Hence, we aimed to evaluate the accuracy of children's school lunch entries in the first Web-based dietary assessment tool for children in Norway, the WebFR, using direct unobtrusive observation as the reference method.

Methods

Participants

All the 4th graders (8-9 years old) from 4 elementary schools in Bærum, the fifth most populated municipality in Norway and a suburb of the capital city, were invited through the schools from September to December 2013. Convenience sampling was used; selected schools were in a short travel distance for the observers and had a highly cooperative school administration. Verbal and written information was given at plenary school meetings and in school classes to parents/guardians and children, respectively. To be included in the study, children needed an Internet access at their home, and their parents/guardians needed access to email. The final sample consisted of 117 of the 196 invited children (59.7%). The study was conducted in accordance with the Declaration of Helsinki. The Regional Ethical Committee in the South East of Norway found the study to fall outside their remit. Approval from the Norwegian Social Science Data Services was obtained, in addition to child assent and written parental consent from all participants.

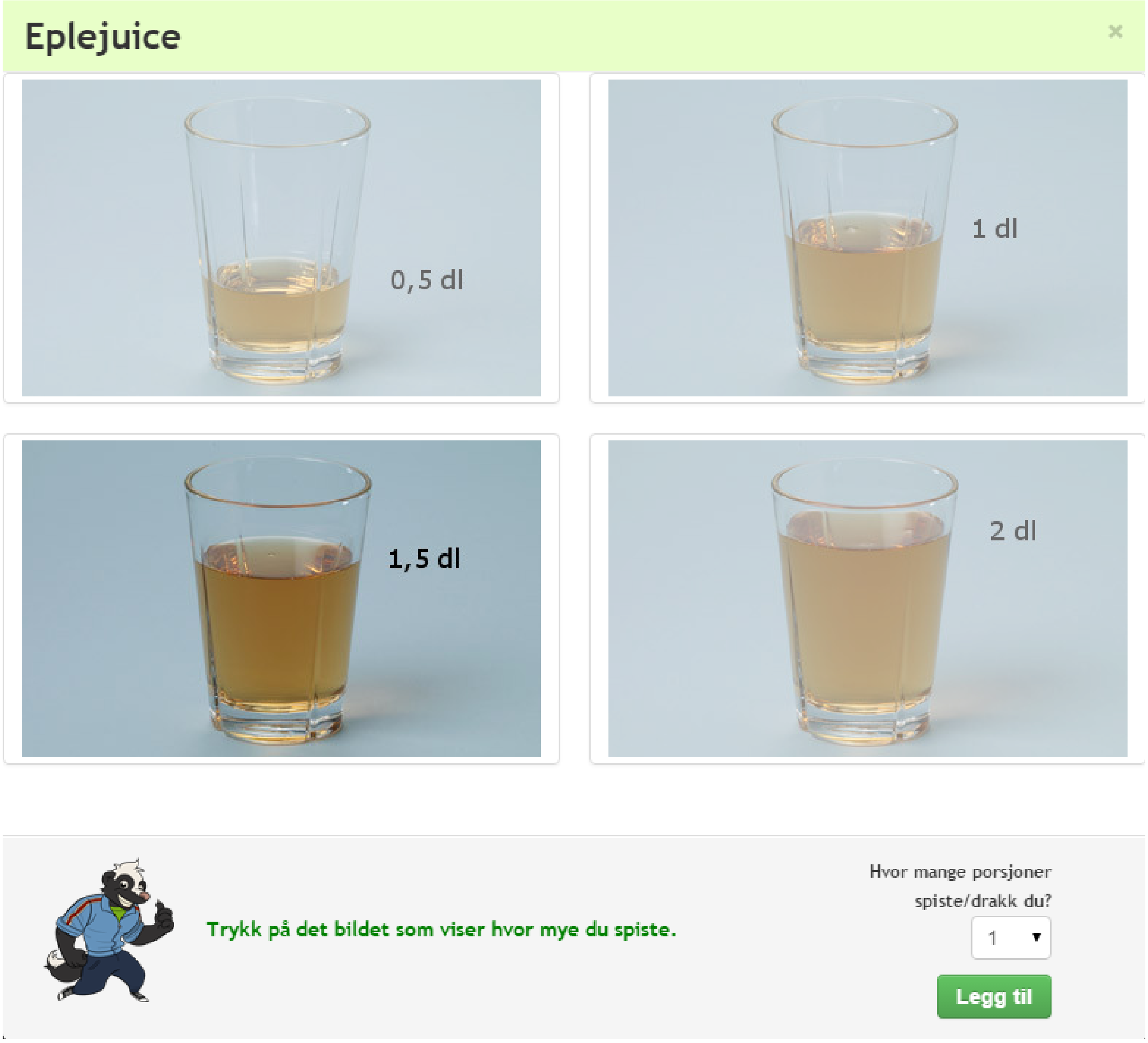
Design

The participants were instructed to record everything they ate and drank in the WebFR, for 4 consecutive days, including a weekend day. They were instructed to complete the recordings in the WebFR at home, with parental assistance, at the end of each recording day, after all meals were consumed. A practical demonstration was given at school in addition to written instructions on how to use the WebFR. During the days they recorded their diet, each child was observed once during school lunch. The children's weights and heights were also measured using standard procedures. After completing the study, the participants received a personal gift card with 2 cinema tickets.

The Web-Based Food Record

The WebFR is designed as a food record, yet including elements of a dietary recall, as recordings are completed by the end of each recording day. It is structured by meals with photos for portion-size assessments. It was adapted from the WebDASC by replacing its food lists with approximately 550 of the most commonly eaten foods and beverages in Norway, based on data from the latest Norwegian National Dietary Survey [17]. In addition, distinctive products designed for children (eg, children's yoghurt) were also included based on Norwegian sales statistics. The photo series included both new photos taken specifically for the purpose of the development of the WebFR and selected photo series from the WebDASC, which were found appropriate to exemplify foods in the WebFR food list. Experienced dietitians considered the appropriateness of all the portion sizes included in the WebFR. All text and audio files were translated and slightly altered. The design of the interface was also changed to some extent; yet, the basic structure and functions of the WebDASC remained untouched. When recording, the participant selects the items consumed for each meal from drop down lists, or by using a free text search field. For each item, the participant chooses the most appropriate picture from a photo series with 2-4 photos illustrating different portion sizes, and then selects the number of portions eaten (Figure 1). Some of the photo series serve as proxies for other food items (eg, a glass of milk is illustrated by a glass of apple juice). When a food item is not found in the WebFR, it is possible to record it in an open field. A voice-assisted cartoon character guides the participant through the recordings. Prompts and pop-ups have been designed to ease recording in the WebFR and counter recall bias. Visuals of the WebFR are shown in Figure 1 and Multimedia Appendix 1.

Figure 1. Screenshot from the Web-based Food Record (WebFR), showing an example of one of the photo series illustrating different portion sizes.



Trykk på det bildet som viser hvor mye du spiste.

Hvor mange porsjoner spiste/drakk du? 1

Observations

The observer team included 1 registered dietician and 2 master's students in nutrition. The observations were performed in classrooms in which the children ate their home-packed lunches during regular school days. Each child was observed one time, during the same period as when they were instructed to record data in the WebFR. Each observer monitored a maximum of 3 children at the same time in an unobtrusive manner (ie, avoiding interaction with the participants and blinding the observations for participants). The children were already familiar with the presence of the observers prior to the observations, through instructional sessions.

The observers used a standardized form to take notes during their observations. To ensure complete recordings, observers were present in the classroom from before the children started eating to until they all had stopped eating. Immediately after each observation session, the observers categorized all observed food items into categories and portion sizes that corresponded to the information in the WebFR, with the aid of tablets containing the lists of categories, items, and all photos found in the WebFR. When the observed foods were not found in the WebFR, the observers described the food item in detail in text and chose the food category and portion size they considered most appropriate for the specific food item. After completion of the data collection, the observer team determined what constituted matches, omissions, and intrusions, using a strict definition; that is, a match was considered a match only when the child and observer clearly described the same item.

Observer training prior to data collection was conducted over a period of 3 weeks, based on the training protocol by Richter et al [18]. Interobserver reliability (IOR), which was expressed as the proportion of agreeing observations between observer pairs, was assessed during the training period and continuously during data collection. The overall agreement of observed food items was 92%, ranging from 88% to 96% between the 3 observer pairs. Lower agreement was found for observed portion sizes, with an overall agreement of 81% and a range from 77% to 88% between observer pairs.

Variables

Variables for “matches,” “omissions,” and “intrusions” were created by comparing the observational data with the participants’ school lunch recordings in the WebFR. Matches are items that are both observed as eaten and recorded as eaten by the child; omissions are items that are observed as eaten but not recorded as eaten; and intrusions are items not observed as eaten, but recorded as eaten by the child.

Participants’ height and weight were measured according to standard procedures, without shoes and in light clothing, to the nearest millimeter and 0.1 kg, respectively, by trained personnel. A digital scale was used (TANITA TBF-300, Tanita Corporation, Tokyo, Japan), in the privacy of a separate room, for each participant. Age and sex-specific body mass index (ISO-BMI) cutoffs defining overweight and obesity among the study participants were applied [19].

Parents/guardians provided information in the written consent form regarding each participant’s sex and age, parental education level (low, intermediate, or high), parental ethnicity (at least one versus no parents/guardians of Norwegian origin), and family structure (mother and father of participant living in same household versus other).

Statistical Analyses

MS Excel (version 2010, Microsoft, Redmond, WA, USA) was used to create all the variables. IBM SPSS (version 21.0, 2012, IBM Corp, New York, NY, USA) was used in all analyses, with the exception of the bias-reduced logistic regression analysis, for which the statistical package R (version 3.0.1, 2013, The R Foundation for Statistical Computing, Vienna, Austria) was used.

Descriptive statistics for the observed food items, recorded food items, matches, omissions, and intrusions were performed. The rates of matches, omissions, and intrusions were calculated for each participant both for all food items combined and at the food item category level (eg, “fruit, berries,” “bread products”). Definitions of these variables are in accordance with previous definitions developed by Baxter et al [20], but without using weighted values, ie, all food items were given equal statistical weight, and thus equal importance, in this study. “Coinciding omissions and intrusions” were also calculated, that is, cases in which a participant has an omission that corresponds to an intrusion within the same food category and within the same meal (eg, “apple” omitted and “pear” intruded during the same school lunch). The portion sizes of the omitted and intruded food items were counted using 4 different categories as follows: extra small (XS), small (S), medium (M), and large (L), based on the available photo series in the WebFR. Unclear observations of food items or portion sizes were excluded from analyses, such as the amount of water consumed from an opaque drinking bottle.

Univariate analyses were conducted to find possible differences in the match rates, omission rates, and intrusion rates as continuous variables, for all foods combined, with regard to the following variables: sex, BMI category, parental educational level, parental ethnicity, and family structure. Parametric tests were used when appropriate. Because the omission rate is the inverse of the match rate (match rate=100 - omission rate), testing for the match rate was therefore equivalent to testing for the omission rate.

A log transformation of the match rate variable was conducted; nevertheless, the assumptions for doing a multivariate linear regression were not present. Hence, match rates were further recoded to a dichotomous variable, which was defined as either a “low match rate” (≤70%) or “high match rate” (>70%). Logistic regression analyses were used to investigate the association between participant characteristics and the quality of the recordings in the WebFR (ie, low versus high match rate). Because of low cell counts, Logistf (bias-reduced logistic regression, Firth correction) [21] was also conducted in the statistical package R to compare the results with those from the logistic regression. The reporting in this study is in line with the STROBE guidelines [22].

Results

The characteristics of the study sample are shown in Table 1. A total of 15 of 117 participants (12.8%) were overweight or obese. Furthermore, the parental education level was high among 77 of 111 children (69.4%), and a great majority had at least one parent/guardian of Norwegian origin.

Table 1. Characteristics of participants (N=117) in a validation study of a Web-based Food Record in Norway.

Characteristics		n	%
Age, years			
	8	13	11.1
	9	104	88.9
Sex			
	Girls	64	54.7
	Boys	53	45.3
ISO-BMI cutoff categories			
	Normal weight	102	87.2
	Overweight or obese	15	12.8
Parental education level [a]			
	Low[b]	12	10.8
	Intermediate[c]	22	19.8
	High[d]	77	69.4
Parental Ethnicity [e]			
	At least one parent/guardian of Norwegian origin	105	91.3
	Both parents/guardians of ethnic origin other than Norwegian	10	8.7
Family Structure [f]			
	Mother and father of participant living in same household	87	78.4
	Other	24	21.6

[a] Information from 111 participants was available for “parental education level.” Complete information on both parents/guardians was available from 108 participants; the 3 cases with missing information from 1 parent/guardian were included in the table based on the 1 available parent/guardian's educational level.

[b] Both parents/guardians' education was maximum high-school level.

[c] One parent/guardian's education was maximum high-school level, and the second parent/guardian's education was at university-college or university level.

[d] Both parents/guardians' education was at the university college or university level.

[e] Information from 115 participants was available for “parental ethnicity.”

[f] Information from 111 participants was available for “family structure.”

Table 2 shows omission rates and intrusion rates for all food items combined, and for categories of food items, listed in descending order from the most to the least frequently observed. The overall averages for the omission and intrusion rates were 27% and 19%, respectively. At the food category level, the average rates varied widely between categories; “bread products” and “milk” were both frequently eaten, and had the lowest omission rates at 5% and 6%, respectively. “Spreads” were eaten most frequently, but the degree of omissions was higher. Food items in the categories “fruit, berries” and “vegetables, salads” were the third and fourth most frequently eaten, also with a high degree of omissions. By contrast, “biscuits, buns, waffles, cakes, and candy” were eaten infrequently, but had the highest proportion of omissions (85%). For intrusions, “bread products” and “dinner leftovers” had the lowest rates, and “beverages, other” and “yogurt” the highest.

Table 2. Omission rate[a] and intrusion rate[b] within different food categories, listed in descending order from the most to the least frequently observed, for all 8- and 9-year old children (N=117) in a validation study of a Web-based Food Record in Norway.

	Omission rate %		Intrusion rate %		Coinciding omissions and intrusions[c]	
	N[d]	Mean (SD)	N[d]	Mean (SD)	N[e]	n (%)
All food items	117	27 (27)	117	19 (26)	136	18 (13.2)
Spreads	93	29 (43)	79	17 (33)	41	7 (17.1)
Bread products	95	5 (22)	97	7 (26)	5	3 (60.0)
Fruit, berries	42	39 (48)	36	25 (44)	22	1 (4.5)
Vegetables, salads	33	45 (49)	23	21 (39)	23	0 (0.0)
Milk	49	6 (24)	52	12 (32)	3	1 (33.3)
Beverages, other[f]	44	18 (39)	62	42 (50)	8	2 (25.0)
Dinner leftovers	17	33 (43)	14	7 (27)	7	0 (0.0)
Miscellaneous	17	44 (50)	12	21 (40)	8	1 (12.5)
Biscuits, buns, waffles, cakes, and candy	12	85 (31)	4	38 (48)	12	1 (8.3)
Yogurt	11	64 (50)	9	56 (53)	7	2 (28.6)

[a] Omission rate = omissions/observed eaten food items × 100 = omissions/(omissions + matches) × 100. Omission rates were calculated for each participant within the different food categories. Participants who were not observed eating foods within a certain category (eg, “fruit, berries”) were excluded from the analyses for this category, regardless of what was recorded eaten.

[b] Intrusion rate = intrusions/recorded eaten food items × 100 = intrusions/(intrusions + matches) × 100. Intrusion rates were calculated for each participant within the different food categories. Participants who did not record eating foods within a certain category (eg, “fruit, berries”) were excluded from the analyses for this category, regardless of what was observed eaten.

[c] Cases where a participant had an omission that corresponds to an intrusion, within the same food category and within the same meal. For example, “apple” omitted and “pear” intruded during the same school lunch. Formula used: coinciding omissions and intrusions/omissions × 100.

[d] Number of participants included in analyses.

[e] Number of food items included in analyses.

[f] Of all intruded “beverages, other” 96% are drinking water.

In addition, Table 2 shows that 18 of all 136 omissions (13.2%) were “coinciding omissions and intrusions”; this proportion varied greatly within the different food item categories. Out of the small number of omissions in the “bread products” category, 3 out of 5 (60%) were coinciding omissions and intrusions, thus most of the omissions in this category were minor errors (eg, white bread replaced by whole grain bread). By contrast, the categories “spreads,” “fruit, berries,” and “vegetables, salads” had low proportions of coinciding omissions and intrusions, and thus most of the omissions in these categories were food items that the participants simply did not record.

Omissions and intrusions of large portion sizes are considered to be more severe than the omission or intrusion of small portion sizes. In Table 3, the proportion of different portion sizes for omissions and intrusions is presented for all food items combined, and for each food category. Although all types of portion sizes were omitted overall, there was a trend toward omitting smaller portions. This was not the case for overall intrusions, in which the medium- and large-portion sizes intruded more often than the smaller ones.

Table 3. Proportion of different sizes of omitted[a] and intruded[b] food items during school lunch for all 8- and 9-year-old participants (N=117) in a validation study of a Web-based Food Record in Norway.

Items	N[d]	Proportion of different sizes[c] of omitted food items, n (%)					N[f]	Proportion of different sizes[c] of intruded food items, n (%)				
		XS	S	M	L	Missing[e]		XS	S	M	L	Missing[e]
All food items	136	28 (20.6)	29 (21.3)	21 (15.4)	22 (16.2)	36 (26.5)	91	9 (9.9)	24 (26.4)	30 (33.3)	28 (30.8)	—
Spreads	41	7 (17.1)	7 (17.1)	12 (29.3)	3 (7.3)	12 (29.3)	22	2 (9.1)	10 (45.5)	6 (27.3)	4 (18.2)	—
Bread products	5	—	1 (20.0)	1 (20.0)	3 (60.0)	0 (0.0)	7	—	0 (0.0)	5 (71.4)	2 (28.6)	—
Fruit, berries	22	10 (45.5)	5 (22.7)	0 (0.0)	3 (13.6)	4 (18.2)	12	2 (16.7)	2 (16.7)	3 (25.0)	5 (41.7)	—
Vegetables, salads	23	5 (21.7)	9 (39.1)	5 (21.7)	1 (4.3)	3 (13.0)	7	1 (14.3)	4 (57.1)	1 (14.3)	1 (14.3)	—
Milk	3	0 (0.0)	0 (0.0)	0 (0.0)	2 (66.7)	1 (33.3)	6	0 (0.0)	0 (0.0)	2 (33.3)	4 (66.7)	—
Beverages, other	8	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	8 (100.0)	26	2 (7.7)	8 (30.8)	9 (34.6)	7 (26.9)	—
Dinner leftovers	7	1 (14.3)	2 (28.6)	2 (28.6)	1 (14. 3)	1 (14.3)	1	1 (100)	0 (0.0)	0 (0.0)	0 (0.0)	—
Miscellaneous	8	1 (12.5)	0 (0.0)	1 (12.5)	2 (25.0)	4 (50.0)	3	0 (0.0)	0 (0.0)	3 (100.0)	0 (0.0)	—
Biscuits, buns, waffles, cakes, and candy	12	4 (33.3)	5 (41.7)	0 (0.0)	2 (16.7)	1 (8.3)	2	1 (50.0)	0 (0.0)	1 (50.0)	0 (0.0)	—
Yogurt	7	—	0 (0.0)	0 (0.0)	5 (71.4)	2 (28.6)	5	—	0 (0.0)	0 (0.0)	5 (100.0)	—

[a] Items observed eaten, but not recorded.

[b] Items recorded, but not observed eaten.

[c] Portion sizes were divided into the following categories: XS=extra small, S=small, M=medium, L=large, based on the photo series available for each food item.

[d] Number of omitted food items included in analyses.

[e] Portion size not possible to observe with certainty, that is, when participants drank from dark-colored drinking bottles or milk cartons, or when participants ate a sandwich where spreads were partially hidden because it was placed in between 2 slices of bread.

[f] Number of intruded food items included in analyses.

The very few omissions in the “bread products” and “milk” categories were mostly of large portion sizes, whereas the omitted portion sizes from “spreads” were mostly of medium sizes. By contrast, the majority of omitted items in the categories “fruit, berries” and “vegetables, salads” were of small portion sizes.

Along the same lines as the omissions, the few intrusions in the categories “bread products” and “milk” were all of medium or large sizes. In the categories “fruit, berries,” “vegetables, salads,” and “spreads,” intrusions occurred for all portion sizes.

Mean rates within subgroups are presented in Table 4. Children with normal weight tended to have lower omission rates, than their overweight or obese peers. The omission rates differed in a statistically significant fashion between the parental education levels; higher omission rates were associated with lower parental educational levels (*P*=.008). Furthermore, we found a statistically significant lower omission rate among children with at least one parent/guardian of Norwegian origin in comparison with children having both parents/guardians of ethnic origins other than Norwegian (*P*=.04). A lower omission rate was also observed among participants from homes in which the mother and father lived together, compared with children from homes with a different family structure.

Table 4. Match rate,[a] omission rate,[b] and intrusion rate[c] within different subgroups among the 8- and 9-year-old participants (N=117) observed during school lunch in a validation study of a Web-based Food Record in Norway.

Variables		Total (N)	Match rate %		Omission rate %		Intrusion rate %	
			Mean (SD)	*P* [d]	Mean (SD)	*P* [d]	Mean (SD)	*P* [d]
Total participants (N)		117	73 (27)		27 (27)		19 (26)	
Sex				.59		.59		.28
	Girls	64	71 (30)		29 (30)		22 (29)	
	Boys	53	76 (22)		24 (22)		16 (23)	
ISO-BMI cutoff categories				.44		.44		.80
	Normal weight	102	74 (27)		26 (27)		19 (26)	
	Overweight or obese	15	69 (27)		31 (27)		21 (28)	
Parental education level [e]				.008		.008		.006
	Low[f]	12	52 (32)		48 (32)		40 (38)	
	Intermediate[g]	22	69 (31)		31 (31)		24 (32)	
	High[h]	77	77 (24)		23 (24)		15 (21)	
Parental ethnicity [i]				.04		.04		.49
	At least one parent/guardian of Norwegian origin	105	75 (26)		25 (26)		19 (26)	
	Both parents/guardians of other ethnic origin than Norwegian	10	57 (28)		44 (28)		24 (27)	
Family structure [j]				.08		.08		.86
	Mother and father of participant living in same household	87	75 (27)		25 (27)		20 (26)	
	Other	24	64 (29)		36 (29)		21 (31)	

[a] Match rate = matches/observed eaten food items × 100 = matches/(omissions + matches) × 100. Match rates were calculated for each participant, for all food items combined.

[b] Omission rate = omissions/observed eaten food items × 100 = omissions/(omissions+ matches) × 100. Omission rates were calculated for each participant, for all food items combined.

[c] Intrusion rate = intrusions/recorded eaten food items × 100 = intrusions/(intrusions+ matches) × 100. Intrusion rates were calculated for each participant, for all food items combined.

[d] *P* value for comparison of groups. Analysis of variance and *t* test were used when applicable; if not, the nonparametric Mann-Whitney or Kruskal-Wallis test was used.

[e] Information from 111 participants was available for "parental education level." Complete information on both parents/guardians was available from 108 participants; the 3 cases with missing information from 1 parent/guardian were included in the table based on the 1 available parent/guardian's educational level.

[f] Both parents/guardians' education was maximum high-school level.

[g] One parent/guardian's education was maximum high-school level, and the second parent/guardian's education was at the university college or university level.

[h] Both parents/guardians' education was at the university college or university level.

[i] Information from 115 participants was available for "parental ethnicity."

[j] Information from 111 participants was available for "family structure."

For intrusion rates, the differences between groups were not statistically significant, except for parental education wherein higher intrusion rates were associated with lower parental educational levels (*P*=.006).

The logistic regression model in Table 5 shows that parental ethnic background and parental education level were the most important variables associated with a "low match rate" (≤70%). Although the "low educational level" effect was reduced when adjusting for other variables, this variable was still borderline significant. Parental ethnicity was the single most important variable associated with a low match rate; if both parents/guardians had an ethnic background other than Norwegian, the odds for a "low match rate" (≤70%) were higher than if at least one parent/guardian was of Norwegian ethnicity. However, the confidence intervals were wide. The results from the Logistf (bias-reduced logistic regression, Firth correction), analyzed due to low cell counts, were consistent with the results from the logistic regression model. Thus, the standard logistic regression was kept as the final model.

Table 5. Variables associated with having a low match rate (≤70%) among 8- and 9-year-old children recording in a Web-based Food Record compared with unobtrusive school lunch observation in Norway.

Variables		n (%) of children		Odds ratio (95% CI)	
		Overall (N=111)	With low match rate (≤70%) (N=44)	Unadjusted (N=111)	Adjusted[a] (N=111)
BMI category [b]					
	Normal weight	96 (86.5)	36 (81.8)	1	1
	Overweight or obese	15 (13.5)	8 (18.2)	1.9 (0.6-5.7)	1.6 (0.4-5.4)
Parental ethnicity [c]					
	Norwegian origin	101 (91.0)	36 (81.8)	1	1
	Non-Norwegians	10 (9.0)	8 (18.2)	7.2 (1.5-35.9)	6.9 (1.3-36.4)
Parental education level					
	High	77 (69.4)	25 (56.8)	1	1
	Intermediate	22 (19.8)	10 (22.7)	1.7 (0.7-4.6)	1.6 (0.6-4.5)
	Low	12 (10.8)	9 (20.5)	6.2 (1.6-25.1)	3.8 (0.9-17.2)
Family structure [d]					
	Mother and father of participant living in same household	87 (78.4)	31 (70.5)	1	1
	Other	24 (21.6)	13 (29.5)	2.1 (0.9-5.3)	2.0 (0.7-5.3)

[a] Adjusted for all other variables in the model in a logistic regression analyses.

[b] ISO-BMI cutoffs applied.

[c] Both parents/guardians of ethnic origin other than Norwegian, compared with at least one parent/guardian of Norwegian origin (reference).

[d] Family structure defined as everything else but “mother and father of participant living in same household” (ie, other) compared with “mother and father of participant living in same household” (reference).

Discussion

Main Findings

We found that 8-9-year-old children on average had a match rate of 73%, an omission rate of 27%, and an intrusion rate of 19%, when comparing parental-assisted entries of school lunch data in a WebFR with unobtrusive observations. Mean omission and intrusion rates for different food categories varied greatly. Lower parental educational levels and a non-Norwegian background were associated with less accurate recordings, but this must be interpreted with caution because of the low numbers in these subgroups.

Comparisons With Other Work

Only a few other validation studies of Web-based 24-hour recalls/records for children have used observation during school meals as a reference method. Among these studies are the one on the Automated Self-Administered 24-hour Recall-Kids-2012 (ASA24-Kids-2012) among 9-11-year olds by Diep et al [23] in the United States, the Food Intake and Physical Activity of School Children (CAAFE) study among 7-10-year olds by Davies et al [12] in Brazil, the Portuguese Self-Administered Computerised 24-h Dietary Recall (PAC24) study among 7-10-year olds by Carvalho et al [24] in Portugal, the WebDASC study among 8-11-year olds by Biltoft-Jensen et al [25] in Denmark, and the study of the modified Self-Administered Children and Infant Nutrition Assessment (SACINA) used among 6-8-year olds by Hunsberger et al [26] in Sweden.

Our results are not directly comparable with these validation studies, partly because the rates of matches, omissions, and intrusions were not calculated in the same way as they were in our study. Nonetheless, we assert that it is possible to interpret the direction of the findings; in the CAAFE and ASA24-Kids-2012 studies, lower agreement between the recordings in the Web-based assessment tool and observations of school lunch were reported than in our study. The CAAFE study had average rates of 44% matches, 30% omissions, and 26% intrusions [12], whereas the ASA24-Kids-2012 study had average rates of 37% matches, 35% omissions, and 27% intrusions [23]. Parental assistance during recordings was encouraged and accomplished for most participants in our study. This was not an option in the CAAFE and ASA24-Kids-2012 studies, although children could ask simple questions to the researchers during recording in these studies. The lack of parental/adult assistance may partly explain differences in the results between studies, and we argue that children at this age need help when recording, which is also proposed elsewhere [8,23,27-29]. Furthermore, a low parental educational level can somewhat explain the dissimilar results between studies; a low educational level was associated with poorer recordings in both the CAAFE study and our study, where 64% and 11% had a low parental educational level, respectively. In addition, in the CAAFE and ASA24-Kids-2012 studies, the children recorded entries after a 24-hour time lag, whereas in our study, the children were instructed to record their data within the same evening. This difference may explain why the match rate was higher in our study because it is demonstrated that omissions and intrusions in children's dietary recalls increase as a function of time [30].

The PAC24 study shows results that are more in line with our results, despite the lack of parental/adult assistance during recording [24]. An explanation could be the broad definition of matches applied in the PAC24 study, in contrast to our study in which matches were defined in a stricter manner.

Because the WebFR is a Norwegian version of the Danish WebDASC, we expected the results to be consistent with the findings from the WebDASC validation study [25]. Surprisingly, Biltoft-Jensen et al [25] found 82% matches, 3% omissions, and 14% intrusions for total foods and beverages, which are remarkably better than in our study, and their rates of omissions were very low in comparison with our 27% omission rate. This large discrepancy cannot be explained by the fact that their calculations were performed slightly differently than those in our study. Age affects children's dietary reports [13,14]. Thus, we argue that age may partly explain the differences between the studies because the children were on average a year older in the WebDASC study than those in this study. In addition, we suggest that reactivity may have been a larger problem in the Danish version than in ours because of their interactive observation style; children were instructed to unpack their packed lunches, separate items, open up sandwiches, and place them on a plate before a photograph was taken. In addition, questions were asked regarding food trading and earlier snacking from their packed lunch [25].

A very high reporting accuracy was reported in the small validation study (n=25) of SACINA by Hunsberger and co-workers; in their study, overall food matches ranged from 86% to 98% [26]. Although children in this study were only to recall 1 lunch meal eaten the previous day, assisted by a dietitian using the Web-based SACINA instrument providing photos and portion estimates, we believe this cannot explain why the accuracy was so contrastingly high compared to other studies.

Baxter et al [20,30-35] compared 9-year-old children's dietary reports in the form of traditional recalls (not Web-based) with direct observations of school meals in several studies. In some of these studies, the rates of omissions and intrusions were calculated and presented in a way that is comparable with our study; the results demonstrate that the omission and intrusion rates varied widely, and that reporting accuracy was reduced when the time intervals between eating and reporting increased. Our results are consistent with or better than the majority of the studies by Baxter et al for "same day recalls."

Only a few studies report the rates of omissions and intrusions for selected food subcategories that are comparable with our findings. Vegetables and sweets were reported as the most often omitted food items in the PAC24 study, whereas beverages were the most commonly intruded item [24]. This is in line with the high omission rates for "vegetables and salads" and "biscuits, buns, waffles, cakes, and candy," and the high intrusion rate for "beverages, other" found in our study. Nevertheless, Biltoft-Jensen et al [25] reported remarkably lower omissions for fruits and vegetables than in our study. Once more, we argue that reactivity may have been a large problem in Biltoft-Jensen et al's study, and may explain the discrepancy between the studies.

To our knowledge, we are the first to report on "coinciding omissions and intrusions," and by doing so we add important knowledge as to whether the omissions and intrusions represent major errors, and not just slightly imprecise recordings. The food category "spreads" had a high omission rate, and most of the omissions were major errors, not "coinciding omissions and intrusions." This discovery has already led us to improve the WebFR, by including tailor-made prompts for "spreads."

Taking the portion sizes of the omitted and intruded food items into account is important because it provides a better understanding of whether these omissions and intrusions are of great concern or not. We observed high omission rates in the food categories "fruit, berries" and "vegetables, salads"; however, the portion sizes of these categories were mostly small in contrast to the portion sizes of omitted "spreads." Thus, we argue that the omissions of spreads are more troublesome than the omissions of fruits and vegetables in our WebFR.

Lower parental education levels have been associated with a higher degree of misreporting among children in the form of underreporting, or both underreporting and overreporting [36-38]. As already described, both our findings and the results of the CAAFE study show that lower parental education was associated with more recording errors [12]. However, the picture is a bit unclear when searching the literature for associations between ethnic backgrounds and omissions or intrusions in dietary assessment tools for children. Baxter et al [30,35] reported that there were no statistically significant differences in rates with regard to ethnicity. Yet, Baranowski et al [39] reported that Hispanic children described more problems when using the Food Intake Recording Software System, an early version of the ASA24-Kids-2012, than other children. Our study is coherent with the latter; although we must emphasize that there were a limited number of participants who had both parents of non-Norwegian origin. A possible explanation for our findings may be that having an ethnicity different from the majority is a proxy for having a different food culture and perhaps literary difficulties, which may be a barrier, even though participants can enter any type of food into the WebFR using the open field option. Furthermore, although audio files were included to assist those with lower reading skills, the WebFR still requires knowledge of the Norwegian language, and thus errors during recording may occur more often among persons with language difficulties. Consequently, we suggest that children and/or parents or guardians with language difficulties should be identified and given extra instructions on how to use the WebFR in future studies. They may benefit from direct personal contact with someone from the research crew, to ensure that they understand what to do.

Studies indicate that underreporting among children increases as BMI increases [37,40,41]. Nevertheless, the odds for a “low match rate” (≤70%) were only slightly higher and not significantly different for overweight/obese than normal weight children in our study. We believe this nonsignificant result may be explained by lack of power.

Strengths and Limitations

The use of direct unobtrusive observations is one of the strengths of this study, because these provide exact information about what is consumed, without affecting the recordings [42,43]. Furthermore, we have demonstrated that our observations of home-packed school lunches were satisfactory; the overall consistency between observers for food items was 92%, which is considered sufficient [15,16,44], and in line with other IOR assessments conducted during observations of home-packed lunches [18]. Lower agreement was seen for the observed portion sizes. The combination of thorough training sessions prior to the study start and IOR assessments before and continuously during the study was important, and it provides a premise for high internal validity.

The small number of individuals in some of the subgroups is a limitation of this study, as the preferable adjustment for cluster effects (school level) proved infeasible due to lack of established statistical methods. Hence, the point estimates in the logistic regression analysis should be interpreted with caution.

For practical reasons, observations were restricted to school lunches and to children in 4th grade (8-9 years). Thus, a limitation is that we do not know whether our findings can be extrapolated to other meals or age groups. In addition, our participants had more highly educated parents/guardians and were less overweight or obese, than the average Norwegian population in which 29% have higher education [45], and 16% of third graders are overweight or obese [46]. However, the proportion of participants with non-ethnic-Norwegian parents/guardians was 20% (see Table 4), which is higher than the 14% average in the general Norwegian population [47]. Despite these differences, this study was performed in a similar setting to the one intended for later use, that is, the next national dietary survey among children in Norway. This contributes to its external validity in a positive manner.

Conclusion

We have demonstrated that 8-9-year-old children had a mean match rate of 73% when recording their food intake from school lunch, with parental assistance, in a WebFR. Some children had difficulties recording, but the mean results were better than what have been reported in most validation studies of other Web-based dietary assessment tools among children. The WebFR could be improved further by including additional prompts for high omission rate foods. We suggest that children and their parents/guardians with language difficulties should be given extra support and information about how to use the WebFR in future studies.

Acknowledgments

The authors thank Susanne Strohmaier for assistance with the statistical analyses.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Selected screenshots from the Web-based Food Record (WebFR).

[WMV File (Windows Media Video), 3MB - jmir_v17i12e273_app1.wmv]

References

1. McPherson R, Hoelscher D, Alexander M, Scanlon K, Serdula M. Dietary assessment methods among school-aged children: Validity and reliability. Prev Med 2000 Aug;31(2):S11-S33. [doi: 10.1006/pmed.2000.0631]
2. Thompson FE, Subar AF. Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi MG, editors. Nutrition in the Prevention and Treatment of Disease. 3rd edition. San Diego, CA: Academic Press; 2008.
3. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc 2010 Jan;110(1):48-51 [FREE Full text] [doi: 10.1016/j.jada.2009.10.008] [Medline: 20102826]
4. Forster H, Fallaize R, Gallagher C, O'Donovan CB, Woolhead C, Walsh MC, et al. Online dietary intake estimation: The Food4Me food frequency questionnaire. J Med Internet Res 2014;16(6):e150 [FREE Full text] [doi: 10.2196/jmir.3105] [Medline: 24911957]
5. Christensen SE, Möller E, Bonn SE, Ploner A, Wright A, Sjölander A, et al. Two new meal- and web-based interactive food frequency questionnaires: Validation of energy and macronutrient intake. J Med Internet Res 2013 Jun;15(6):e109 [FREE Full text] [doi: 10.2196/jmir.2458] [Medline: 23739995]
6. Adamson AJ, Baranowski T. Developing technological solutions for dietary assessment in children and young people. J Hum Nutr Diet 2014 Jan;27 Suppl 1:1-4. [doi: 10.1111/jhn.12206] [Medline: 24450502]

7. Biltoft-Jensen A, Trolle E, Christensen T, Islam N, Andersen LF, Egenfeldt-Nielsen S, et al. WebDASC: A web-based dietary assessment software for 8-11-year-old Danish children. J Hum Nutr Diet 2014 Jan;27 Suppl 1:43-53. [doi: 10.1111/j.1365-277X.2012.01257.x] [Medline: 22594587]
8. Baranowski T, Islam N, Douglass D, Dadabhoy H, Beltran A, Baranowski J, et al. Food Intake Recording Software System, version 4 (FIRSSt4): A self-completed 24-h dietary recall for children. J Hum Nutr Diet 2014 Jan;27 Suppl 1:66-71 [FREE Full text] [doi: 10.1111/j.1365-277X.2012.01251.x] [Medline: 22616645]
9. Vereecken C, Covents M, Maes L, Moyson T. Formative evaluation of the dietary assessment component of Children's and Adolescents' Nutrition Assessment and Advice on the Web (CANAA-W). J Hum Nutr Diet 2014 Jan;27 Suppl 1:54-65. [doi: 10.1111/j.1365-277X.2012.01290.x] [Medline: 23980931]
10. Foster E, Hawkins A, Delve J, Adamson AJ. Reducing the cost of dietary assessment: Self-completed recall and analysis of nutrition for use with children (SCRAN24). J Hum Nutr Diet 2014 Jan;27 Suppl 1:26-35. [doi: 10.1111/jhn.12108] [Medline: 23662627]
11. Moore HJ, Hillier FC, Batterham AM, Ells LJ, Summerbell CD. Technology-based dietary assessment: Development of the Synchronised Nutrition and Activity Program (SNAP). J Hum Nutr Diet 2014 Jan;27 Suppl 1:36-42. [doi: 10.1111/j.1365-277X.2012.01295.x] [Medline: 23848949]
12. Davies VF, Kupek E, de Assis MA, Natal S, Di Pietro PF, Baranowski T. Validation of a web-based questionnaire to assess the dietary intake of Brazilian children aged 7-10 years. J Hum Nutr Diet 2015 Jan;28 Suppl 1:93-102. [doi: 10.1111/jhn.12262] [Medline: 25139011]
13. Livingstone MB, Robson PJ, Wallace JM. Issues in dietary intake assessment of children and adolescents. Br J Nutr 2004 Oct;92 Suppl 2:S213-S222. [Medline: 15522159]
14. Livingstone MB, Robson PJ. Measurement of dietary intake in children. Proc Nutr Soc 2000 May;59(2):279-293. [Medline: 10946797]
15. Simons-Morton BG, Baranowski T. Observation in assessment of children's dietary practices. J Sch Health 1991 May;61(5):204-207. [Medline: 1943043]
16. Baglio ML, Baxter SD, Guinn CH, Thompson WO, Shaffer NM, Frye Francesca HA. Assessment of interobserver reliability in nutrition studies that use direct observation of school meals. J Am Diet Assoc 2004 Sep;104(9):1385-1392 [FREE Full text] [doi: 10.1016/j.jada.2004.06.019] [Medline: 15354155]
17. The Norwegian Directorate of Health. Norkost 3. A nationwide dietary survey among 18-70 year old men and women in Norway-11 [In Norwegian]. Oslo, Norway: The Norwegian Directorate of Health; 2012. URL: http://www.lhl.no/Documents/norkost-3-is-2000%5b1%5d.pdf [accessed 2015-10-07] [WebCite Cache ID 6ajFZ1N8Q]
18. Richter SL, Vandervet LM, Macaskill LA, Salvadori MI, Seabrook JA, Dworatzek PD. Accuracy and reliability of direct observations of home-packed lunches in elementary schools by trained nutrition students. J Acad Nutr Diet 2012 Oct;112(10):1603-1607. [doi: 10.1016/j.jand.2012.07.010] [Medline: 23017569]
19. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: International survey. BMJ 2000 May 6;320(7244):1240-1243 [FREE Full text] [Medline: 10797032]
20. Baxter SD, Thompson WO, Litaker MS, Frye Francesca H A, Guinn CH. Low accuracy and low consistency of fourth-graders' school breakfast and school lunch recalls. J Am Diet Assoc 2002 Mar;102(3):386-395 [FREE Full text] [Medline: 11905461]
21. Firth D. Bias reduction of maximum likelihood estimates. Biometrika 1993;80(1):27-38. [doi: 10.1093/biomet/80.1.27]
22. von EE, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. Lancet 2007 Oct 20;370(9596):1453-1457. [doi: 10.1016/S0140-6736(07)61602-X] [Medline: 18064739]
23. Diep CS, Hingle M, Chen T, Dadabhoy HR, Beltran A, Baranowski J, et al. The Automated Self-Administered 24-Hour Dietary Recall for Children, 2012 Version, for Youth Aged 9 to 11 Years: A validation study. J Acad Nutr Diet 2015 Oct;115(10):1591-1598. [doi: 10.1016/j.jand.2015.02.021] [Medline: 25887784]
24. Carvalho MA, Baranowski T, Foster E, Santos O, Cardoso B, Rito A, et al. Validation of the Portuguese self-administered computerised 24-hour dietary recall among second-, third- and fourth-grade children. J Hum Nutr Diet 2015 Dec;28(6):666-674. [doi: 10.1111/jhn.12280] [Medline: 25420921]
25. Biltoft-Jensen A, Bysted A, Trolle E, Christensen T, Knuthsen P, Damsgaard CT, et al. Evaluation of Web-based Dietary Assessment Software for Children: Comparing reported fruit, juice and vegetable intakes with plasma carotenoid concentration and school lunch observations. Br J Nutr 2013 Jul 14;110(1):186-195. [doi: 10.1017/S0007114512004746] [Medline: 23181984]
26. Hunsberger M, Pena P, Lissner L, Grafström L, Vanaelst B, Börnhorst C, et al. Validity of self-reported lunch recalls in Swedish school children aged 6-8 years. Nutr J 2013;12:129 [FREE Full text] [doi: 10.1186/1475-2891-12-129] [Medline: 24047239]
27. Foster E, Hawkins A, Simpson E, Adamson AJ. Developing an interactive portion size assessment system (IPSAS) for use with children. J Hum Nutr Diet 2014 Jan;27 Suppl 1:18-25. [doi: 10.1111/jhn.12127] [Medline: 23682796]
28. Baranowski T, Islam N, Baranowski J, Martin S, Beltran A, Dadabhoy H, et al. Comparison of a web-based versus traditional diet recall among children. J Acad Nutr Diet 2012 Apr;112(4):527-532 [FREE Full text] [doi: 10.1016/j.jada.2011.10.002] [Medline: 22717216]

29. Burrows TL, Martin RJ, Collins CE. A systematic review of the validity of dietary assessment methods in children when compared with the method of doubly labeled water. J Am Diet Assoc 2010 Oct;110(10):1501-1510. [doi: 10.1016/j.jada.2010.07.008] [Medline: 20869489]
30. Baxter SD, Thompson WO, Davis HC, Johnson MH. Impact of gender, ethnicity, meal component, and time interval between eating and reporting on accuracy of fourth-graders' self-reports of school lunch. J Am Diet Assoc 1997 Nov;97(11):1293-1298. [doi: 10.1016/S0002-8223(97)00309-X] [Medline: 9366868]
31. Baxter SD, Smith AF, Litaker MS, Guinn CH, Shaffer NM, Baglio ML, et al. Recency affects reporting accuracy of children's dietary recalls. Ann Epidemiol 2004 Jul;14(6):385-390. [doi: 10.1016/j.annepidem.2003.07.003] [Medline: 15246326]
32. Baxter SD, Thompson WO, Litaker MS, Guinn CH, Frye Francesca HA, Baglio ML, et al. Accuracy of fourth-graders' dietary recalls of school breakfast and school lunch validated with observations: In-person versus telephone interviews. J Nutr Educ Behav 2003;35(3):124-134 [FREE Full text] [Medline: 12773283]
33. Baxter SD, Guinn CH, Royer JA, Hardin JW, Mackelprang AJ, Smith AF. Accuracy of children's school-breakfast reports and school-lunch reports (in 24-h dietary recalls) differs by retention interval. Eur J Clin Nutr 2009 Dec;63(12):1394-1403 [FREE Full text] [doi: 10.1038/ejcn.2009.107] [Medline: 19756033]
34. Baxter SD, Smith AF, Guinn CH, Thompson WO, Litaker MS, Baglio ML, et al. Interview format influences the accuracy of children's dietary recalls validated with observations. Nutr Res 2003 Nov;23(11):1537-1546. [doi: 10.1016/S0271-5317(03)00179-9]
35. Baxter SD, Thompson WO, Smith AF, Litaker MS, Yin Z, Frye Francesca HA, et al. Reverse versus forward order reporting and the accuracy of fourth-graders' recalls of school breakfast and school lunch. Prev Med 2003 May;36(5):601-614. [Medline: 12689806]
36. Lioret S, Touvier M, Balin M, Huybrechts I, Dubuisson C, Dufour A, et al. Characteristics of energy under-reporting in children and adolescents. Br J Nutr 2011 Jun;105(11):1671-1680. [doi: 10.1017/S0007114510005465] [Medline: 21262062]
37. Rangan AM, Flood VM, Gill TP. Misreporting of energy intake in the 2007 Australian Children's Survey: Identification, characteristics and impact of misreporters. Nutrients 2011 Feb;3(2):186-199 [FREE Full text] [doi: 10.3390/nu3020186] [Medline: 22254091]
38. Murakami K, Miyake Y, Sasaki S, Tanaka K, Arakawa M. Characteristics of under- and over-reporters of energy intake among Japanese children and adolescents: The Ryukyus Child Health Study. Nutrition 2012 May;28(5):532-538. [doi: 10.1016/j.nut.2011.08.011] [Medline: 22113067]
39. Baranowski T, Islam N, Baranowski J, Cullen KW, Myres D, Marsh T, et al. The food intake recording software system is valid among fourth-grade children. J Am Diet Assoc 2002 Mar;102(3):380-385. [Medline: 11902371]
40. Baxter SD, Smith AF, Litaker MS, Guinn CH, Nichols MD, Miller PH, et al. Body mass index, sex, interview protocol, and children's accuracy for reporting kilocalories observed eaten at school meals. J Am Diet Assoc 2006 Oct;106(10):1656-1662 [FREE Full text] [doi: 10.1016/j.jada.2006.07.013] [Medline: 17000199]
41. Fisher JO, Johnson RK, Lindquist C, Birch LL, Goran MI. Influence of body composition on the accuracy of reported energy intake in children. Obes Res 2000 Nov;8(8):597-603. [doi: 10.1038/oby.2000.77] [Medline: 11156436]
42. Baxter SD, Hardin JW, Smith AF, Royer JA, Guinn CH, Mackelprang AJ. Twenty-four hour dietary recalls by fourth-grade children were not influenced by observations of school meals. J Clin Epidemiol 2009 Aug;62(8):878-885 [FREE Full text] [doi: 10.1016/j.jclinepi.2008.11.007] [Medline: 19230605]
43. Smith AF, Baxter SD, Hardin JW, Guinn CH, Royer JA, Litaker MS. Validation-study conclusions from dietary reports by fourth-grade children observed eating school meals are generalisable to dietary reports by comparable children not observed. Public Health Nutr 2007 Oct;10(10):1057-1066 [FREE Full text] [doi: 10.1017/S1368980007683888] [Medline: 17381934]
44. Simons-Morton BG, Forthofer R, Huang IW, Baranowski T, Reed DB, Fleishman R. Reliability of direct observation of schoolchildren's consumption of bag lunches. J Am Diet Assoc 1992 Feb;92(2):219-221. [Medline: 1737906]
45. Statistics Norway. Population's Level of Education. Oslo, Norway: Statistics Norway; 2013 Oct 1. URL: http://www.ssb.no/en/utdanning/statistikker/utniv/ [accessed 2015-10-07] [WebCite Cache ID 6ajFLuQRa]
46. Norwegian Institute of Public Health. Division of Epidemiology. Child Growth in Norway 2008-2010-2012. Height, Weight and Waist Circumference among 3rd Graders [In Norwegian]. Oslo, Norway: Norwegian Institute Of Public Health; 2014. URL: http://www.fhi.no/dokumenter/b8fd82d8a0.pdf [accessed 2015-10-07] [WebCite Cache ID 6ajFwWlvB]
47. Statistics Norway. Immigrants and Norwegian-Born to Immigrant Parents in 13 Municipalities [In Norwegian]. Oslo, Norway: Statistics Norway; 2014. URL: https://www.ssb.no/befolkning/artikler-og-publikasjoner/_attachment/190485?_ts=147aaa8bd38 [accessed 2015-10-07] [WebCite Cache ID 6ajGDso2u]

Abbreviations

ASA24-Kids-2012: Automated Self-Administered 24-hour Recall-Kids-2012
BMI: body mass index
CAAFE: the Food Intake and Physical Activity of School Children
IOR: interobserver reliability

ISO-BMI: age and sex-specific body mass index
PAC24: the Portuguese Self-Administered Computerised 24-h Dietary Recall
SACINA: Self-Administered Children and Infant Nutrition Assessment
WebDASC: Web-based Dietary Assessment Software for Children
WebFR: Web-based Food Record

Edited by G Eysenbach; submitted 13.08.15; peer-reviewed by Å Svensson, V Fernandes Davies; comments to author 03.09.15; revised version received 15.09.15; accepted 02.10.15; published 07.12.15.

Please cite as:
Medin AC, Astrup H, Kåsin BM, Andersen LF
Evaluation of a Web-Based Food Record for Children Using Direct Unobtrusive Lunch Observations: A Validation Study
J Med Internet Res 2015;17(12):e273
URL: http://www.jmir.org/2015/12/e273/
doi:10.2196/jmir.5031
PMID:26680744

Original Paper

Potential and Challenges in Collecting Social and Behavioral Data on Adolescent Alcohol Norms: Comparing Respondent-Driven Sampling and Web-Based Respondent-Driven Sampling

Janina Hildebrand[1], BSc; Sharyn Burns[1], BEd, DipPhyHlthEd, PGDip(HlthPr), MPH, PhD; Yun Zhao[2], BSc, PGDip, MSc, PhD; Roanna Lobo[1], BA Hons, MSc, PhD; Peter Howat[1], DipTchg, DipEd, DPE, MSc, PhD; Steve Allsop[3], BSc Hons, PGDip, PhD; Bruce Maycock[2], PhD, MEd, PGrad Dip (Bus), BPE

[1]Collaboration for Evidence, Research and Impact in Public Health, School of Public Health, Curtin University, Bentley, Australia
[2]School of Public Health, Curtin University, Bentley, Australia
[3]National Drug Research Institute, Curtin University, Perth, Australia

Corresponding Author:
Bruce Maycock, PhD, MEd, PGrad Dip (Bus), BPE
Collaboration for Evidence, Research and Impact in Public Health
School of Public Health
Curtin University
GPO Box U1987
Bentley, 6845
Australia
Phone: 61 9266 7988
Fax: 61 9266 2958
Email: b.maycock@curtin.edu.au

Abstract

Background: Respondent-driven sampling (RDS) is a method successfully used to research hard-to-access populations. Few studies have explored the use of the Internet and social media with RDS, known as Web-based RDS (WebRDS). This study explored the use of combining both “traditional” RDS and WebRDS to examine the influences on adolescent alcohol use.

Objective: This paper reports on the recruitment processes and the challenges and enablers of both RDS and WebRDS. It details comparative recruitment data and provides a summary of the utility of both methods for recruiting adolescents to participate in an online survey investigating youth alcohol norms.

Methods: Process evaluation data collected from research staff throughout the study were used to assess the challenges and solutions of RDS and WebRDS. Pearson chi-square test (Fisher’s exact test if applicable) was used to compare the differences in sociodemographics and drinking behavior between data collected by RDS and WebRDS.

Results: Of the total sample (N=1012), 232 adolescents were recruited by RDS and 780 by WebRDS. A significantly larger proportion of Aboriginal or Torres Strait Islander (*P*<.001) participants who spoke English as their main language at home (*P*=.03), and of middle and lower socioeconomic status (*P*<.001) was found in the RDS sample. The RDS sample was also found to have a higher occurrence of past 7-day drinking (*P*<.001) and past 7-day risky drinking (*P*=.004). No significant differences in gender, age, past month alcohol use, and lifetime alcohol use were observed between the RDS and WebRDS samples. This study revealed RDS and WebRDS used similar lengths of chains for recruiting participants; however, WebRDS conducted a faster rate of recruitment at a lower average cost per participant compared to RDS.

Conclusions: Using WebRDS resulted in significant improvements in the recruitment rate and was a more effective and efficient use of resources than the traditional RDS method. However, WebRDS resulted in partially different sample characteristics to traditional RDS. This potential effect should be considered when selecting the most appropriate data collection method.

(J Med Internet Res 2015;17(12):e285) doi:10.2196/jmir.4762

KEYWORDS

Internet; respondent-driven sampling (RDS); WebRDS; adolescent; alcohol; social media; participant recruitment

Introduction

Research has shown that respondent-driven sampling (RDS) is a viable method to recruit individuals from hard-to-access populations (eg, drug users and sex workers) for which no sampling frame exists [1-8]. It has also been used successfully to examine young people's risk-taking behavior [9-12] in several countries [9,10,13,14] and settings [11,12,14]. RDS is a probability sampling method, which is a modified form of chain referral sampling [15]. RDS uses chain referrals with structured incentives as a recruitment strategy, whereby social network parameters of participants are used to weigh the data to statistically adjust for potential chain referral bias [16]. Samples generated by RDS are generally more heterogeneous than those recruited via other sampling means, thus are potentially more generalizable to the population of interest [8,17,18]. In addition, RDS allows validity and reliability of results and randomization of the sample [13]. RDS is not constrained by certain biases associated with other snowball-type recruitment methods, such as time-space for which the sample is restricted to individuals present in public venues [15,19]. Nonprobability snowball sampling is a nonrandom convenience method from which biased estimation is likely to occur [20]. In contrast, the aforementioned characteristics of RDS enable better representation of social networks and consequently more valid calculation of population estimates [16]. Limitations of RDS include the potentially high cost of data collection, requiring substantial resources in terms of personnel and cost [21]. Zablotska et al [22], in their comparison of data collection methods, concluded that their RDS sample was the most consistent to population estimates, but that it was complex and logistically demanding compared to time-location and online recruitment, which were more cost-effective and easier to implement. RDS estimates have also shown a larger variance compared to simple random sampling [23]. For these reasons, Goel and Salganik [23] suggested RDS may not be an optimal strategy for public health surveillance.

Recently, RDS has adopted Web-based recruitment methods (WebRDS) [21,24]. Two studies reported that WebRDS may be a feasible alternative to traditional RDS in the recruitment of young adults [21,24]. Bauermeister et al [21] used WebRDS to administer an online survey designed to assess alcohol and other drug use among those aged 18 to 24 years in the United States. Study primary participants (seeds) were recruited online through a targeted Facebook advertisement. Compared with traditional RDS or face-to-face recruitment, WebRDS demonstrated an ability to overcome temporal and physical barriers to recruitment by allowing young adults to refer a peer using features on social networking sites, such as a wall post, status update, or personal message. Offering multiple approaches to peer referral was found to be most effective in maximizing rate of data collection and length of referral chains, the total number of waves recruited by a seed. Limitations to achieving a representative sample in this study included racial/ethnic and socioeconomic disparities in computer access and frequency of use. Particularly, low education levels and self-identified nonwhite individuals were underrepresented in the sample [21]. Wejnert and Heckathorn [24] also explored the use of WebRDS to obtain a sample of 159 American undergraduate students, initiated from 9 seeds recruited via the Internet. Their findings support the potential for WebRDS to generate electronic network chains with minimal resources and at a significantly faster rate than traditional RDS methodology. Their study also recommended using small incentives due to the reduced respondent burden associated with WebRDS. Although preliminary results indicate WebRDS may be a more efficient recruitment method compared with traditional RDS and standard sampling strategies [25], there is a dearth of studies using this method in an adolescent population.

This paper describes the adaptive and iterative methods from a study using RDS and WebRDS to recruit a sample of youth aged 14 to 17 years exploring alcohol-related norms and behaviors using an online survey. This paper details the recruitment processes and the challenges and enablers of both RDS and WebRDS and compares recruitment data and the utility of both methods for recruiting and collecting data from adolescents.

Methods

Over a 21-week period, a combination of RDS and WebRDS was used to recruit adolescents aged 14 to 17 years living in Perth, Western Australia, to complete an online survey investigating their alcohol-related social norms and behaviors [26]. The comprehensive survey instrument was substantially longer than those of other WebRDS studies [21,24] and, therefore, had higher potential risk of participant attrition. Challenges and solutions of each recruitment method are discussed subsequently.

Initial study participants using an RDS approach (seeds) were purposively selected from the community, youth programs, and sports clubs. Study eligibility was confirmed face-to-face by a trained staff member through a series of screening questions to determine age (14-17 years), location (Perth metropolitan area), and previous study involvement. The use of WebRDS evolved as a response to the barriers experienced during the traditional RDS recruitment.

Procedure

RDS was conducted face-to-face with participants meeting with a staff member to complete the survey on a tablet device (iPad) or a paper version. The seeds, and their subsequent referrals, were given referral coupons and asked to recruit up to 3 of their peers into the study within 2 weeks of their survey completion. Participants received an AU $15 gift card to an electronics store for completing the survey. For each subsequent successful referral, they received an additional AU $10 gift card for a possible total incentive of AU $45. To meet ethics requirements, each participant completed a mature minor assessment and provided verbal and written informed consent to participate [27]. Participants could cease involvement at any point. See Hildebrand et al [26,28] for full descriptions of RDS methodology and this study's survey procedure.

The mature minor assessment was undertaken by providing potential participants with a verbal and written explanation of the study, explaining the purpose of the study, who was conducting it and the anticipated outcomes, and the requirements of participation. Participant information sheets were pilot-tested with members of the target group and assessed for readability to ensure a reading age of 11 years (grade 6). It was explicitly stated that the decision to participate in the research was that of the young person's alone and that they could withdraw at any time without prejudice. Adolescents were given the opportunity to ask questions along with sufficient time to make a decision. A protocol checklist to establish each adolescent's ability to understand the requirements of the study and their mature minor status before deciding to participate was developed. Using the protocol checklist, cognitive interviewing was adapted and applied to evaluate each participant's comprehension of the study information and consent materials. Potential participants were asked a set of 5 questions pertaining to the study procedures, which they had to answer verbally. Participants who provided verbal response to the questions that demonstrated their understanding of the ethical considerations and participant rights, were assigned mature minor status and were considered able to provide consent and continue participation. Data collection officers completed training, which included explanation of the process, the importance of consistency, opportunity to practice the procedure, and a manual outlining the steps and a set of required responses. The project manager met with the data collection officers on a weekly basis to monitor and collect feedback on the progress [27].

Challenges and Solutions of Respondent-Driven Sampling

Throughout the data collection, we faced a number of challenges due to the age of the target population and implications of adhering to the specific RDS methodology and ethical considerations [27]. The greatest challenge encountered was lack of use of the referral coupons. Only 2 participants initiated contact with the study staff from these coupons. Forty-three participants who had conducted the survey face-to-face were contacted. Feedback on attempts to distribute referral coupons indicated that many participants had not attempted to recruit peers. Some reasons cited for not distributing coupons were forgetting, loss of the coupon, going on holiday, all friends being involved in the initial recruitment, being too busy to pass on, friends being too busy to be involved, and thinking friends would not be interested. Overall, reliance on paper coupons to facilitate referrals did not result in an adequate referral rate or chain length.

Additional resources were allocated to enhance recruitment, including presentations conducted at the beginning of sports training or youth program sessions and increased focus on locations to administer surveys “on the spot” where groups of youth congregated (shopping centers, movie theaters, and the central business district). Referrals were sought and immediately recruited if they were part of the existing group by asking participants if they would invite any of their peers who were present as their referrals. The best opportunities to approach and recruit participants were found to be when youth were spending time with friends, such as after finishing lunch in a food court or while waiting outside a theater for a movie to start, and not engaged in other activities.

Participants recruited from youth programs posed a particular challenge for data collection. Feedback from youth program staff and data collection officers suggest issues such as low literacy levels, participants skipping questions, or not having access to Internet or a personal email account to enable follow-up made data collection difficult. Although these challenges were encountered generally among youth, they were particularly common among at-risk youth. Even when a mobile phone number was provided, there were issues associated with participants running out of phone credit. Participants from youth group programs also commonly claimed to have no relevant peers to refer.

Web-Based Respondent-Driven Sampling

During recruitment, a large proportion of participants indicated that their preferred method of communication was Facebook. As a response to these barriers, WebRDS recruitment was introduced. Following approval by the Curtin University Human Ethics Research Committee, the study protocol was amended to recruit participants online in parallel to the continued face-to-face recruitment. Adolescents who expressed interest via Facebook or in response to seeing a flyer or Facebook advertisement had their eligibility to participate confirmed by a research team member before provision of the survey link and a password. The brief screening questionnaire for eligibility and the mature minor assessment were also incorporated into the online survey in multiple-choice format. This process was adapted for the electronic assessment by providing the study information at the start of the survey and subsequently assessing participants' understanding of the information by asking the same questions as the face-to-face version using a multi-choice format. All questions had to be answered correctly to allow study participation. The process was initially tested face-to-face to ensure its validity.

Contact details of participants were collected at the end of the survey to enable communication with participants for referral purposes and sending incentives in the form of electronic gift cards. Survey responses were checked and transferred to the study database daily and contact details were checked against other participants for duplication. Referral messages were sent to the participant within 24 hours and incentives within 48 hours of survey completion. Referral coupons were provided to the participant via the survey link accompanied by 3 unique identifying passwords within the referral messages. Participants were asked to forward the survey link and passwords to their peers via their preferred social media. The passwords were identical to the referral codes. Passwords expired after use to ensure each code was used only once and to distinguish between participants.

Recruiting via Facebook

One week after the online referral method commenced, a Facebook business page was created for the study to enhance communication with participants and recruitment. Data continued to be collected via a separate survey Web-hosting site. To increase response rates, we also explored the use of Facebook advertisements and post promotions. The advertising campaign ran for 35 days during which 3 different advertisements targeting Facebook users who lived in Perth aged between 14 and 17 years were piloted. Two of the advertisements used keywords or specification of interest groups (eg, drinking, alcohol, binge drinking, drinking culture, alcohol intoxication, sports). Example advertisements consisted of the following text: “Curtin University Study—Earn up to $45 JB Hi-Fi gift cards by completing the Youth Alcohol Norms online survey.”

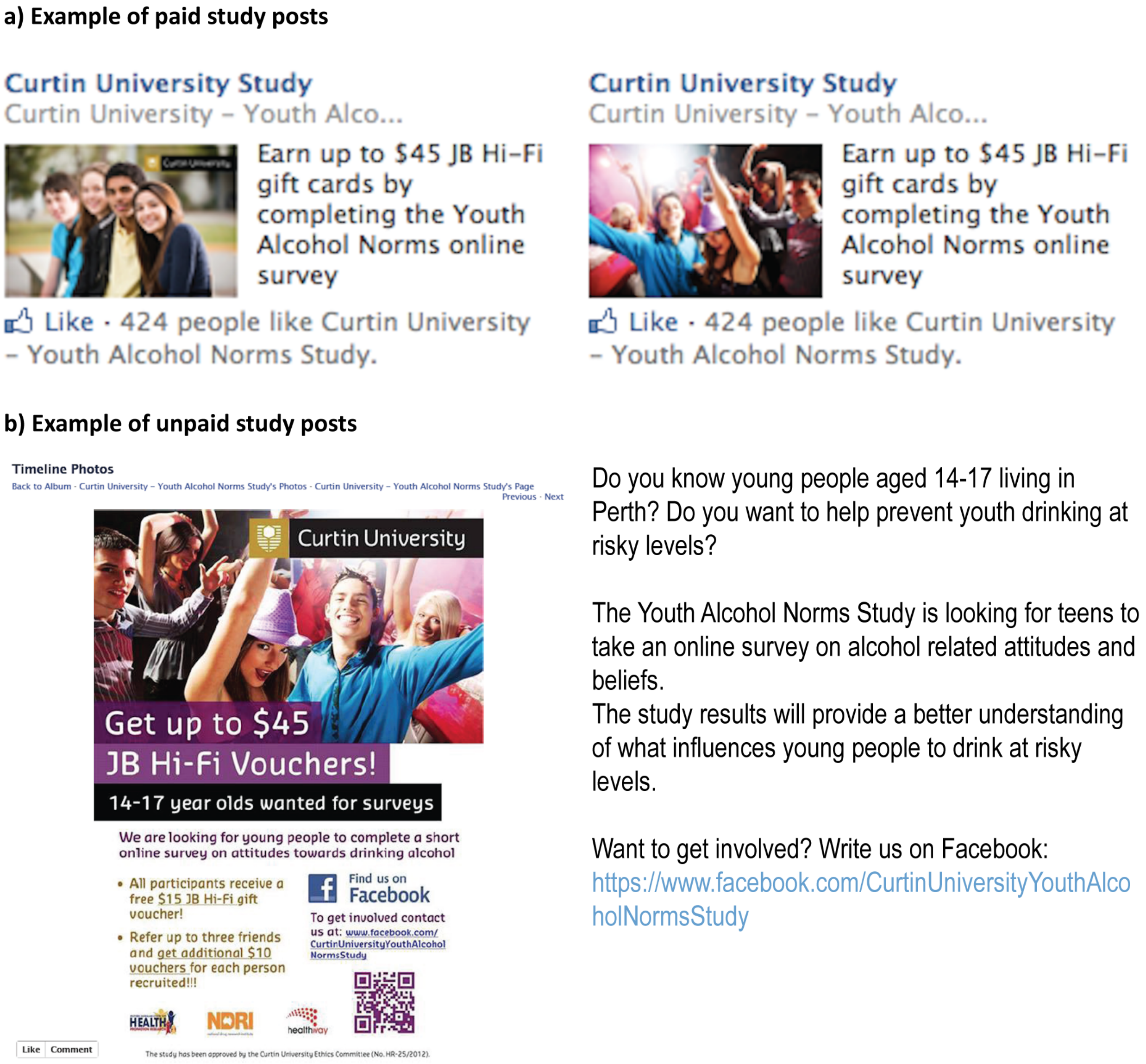
Facebook advertisements were not visible to smartphone and tablet users; hence, page post promotions were also used to boost the visibility and reach of status updates. Status updates are messages/images screened on a Facebook user’s news feed, the center column of an individual’s profile where stories are constantly updated from people and pages that they “like” and then follow. A higher rate of expressions of interest was noted during the time when posts were promoted compared to as a response to the advertisements. Consequently, status updates were used as the main form of promotion.

In total, 6 posts were promoted with different picture and text options. A new post was usually created and promoted on a Friday and/or Saturday afternoon and ran for 24 hours. Facebook page metrics indicated that fans (users who liked our page) were most frequently online Fridays to Sundays, increasing the chances of posts being seen. Five of 6 posts targeted adolescents living in Australia who were between the ages of 14 and 17 years with one of these specifically targeted at youth aged 14 and 15 years to increase participation among this group.

General unpaid status updates were also posted weekly on our page timeline (n=11) as a means of engaging with participants to remind them to complete their survey or invite their peers to the study. To enhance recruitment, several youth-focused sports clubs and youth groups agreed to post the study’s recruitment poster on their Facebook page.

An example text of a status update included: “Are you aged between 14-17? If YES, then we want YOU for our study! You can earn up to $45 JB Hi-Fi gift cards by simply completing the Youth Alcohol Norms Survey and referring your peers!! Send us a private message to get involved!” An example of 2 images and text typical of a study advertisement and unpaid post used during this campaign are shown in Figure 1.

Figure 1. Example of (A) image and text used in study advertisement and (B) unpaid post used during the Facebook campaign.



a) Example of paid study posts

Curtin University Study
Curtin University – Youth Alco...
Earn up to $45 JB Hi-Fi gift cards by completing the Youth Alcohol Norms online survey
Like · 424 people like Curtin University – Youth Alcohol Norms Study.

Curtin University Study
Curtin University – Youth Alco...
Earn up to $45 JB Hi-Fi gift cards by completing the Youth Alcohol Norms online survey
Like · 424 people like Curtin University – Youth Alcohol Norms Study.

b) Example of unpaid study posts

Timeline Photos
Back to Album · Curtin University – Youth Alcohol Norms Study's Photos · Curtin University – Youth Alcohol Norms Study's Page
Previous · Next

Curtin University
Get up to $45
JB Hi-Fi Vouchers!
14-17 year olds wanted for surveys
We are looking for young people to complete a short online survey on attitudes towards drinking alcohol
- All participants receive a free $15 JB Hi-Fi gift voucher!
- Refer up to three friends and get additional $10 vouchers for each person recruited!!!

Find us on Facebook
To get involved contact us at: www.facebook.com/CurtinUniversityYouthAlcoholNormsStudy

The study has been approved by the Curtin University Ethics Committee (No. HR-25/2012).

Do you know young people aged 14-17 living in Perth? Do you want to help prevent youth drinking at risky levels?

The Youth Alcohol Norms Study is looking for teens to take an online survey on alcohol related attitudes and beliefs.
The study results will provide a better understanding of what influences young people to drink at risky levels.

Want to get involved? Write us on Facebook:
https://www.facebook.com/CurtinUniversityYouthAlcoholNormsStudy

Challenges and Solutions of Web-Based Respondent-Driven Sampling

Every effort was made to ensure duplicate survey completion was detected. Participants were initially screened for study eligibility before being able to proceed to the survey. Rapid referrals and similar or identical email addresses and contact numbers to previous participants were investigated. Geodata information about the geographic location of survey completions aided this process to identify respondents who did not meet location criteria. When screening raised questions of eligibility, referral messages and incentives were not sent and participants were contacted by phone, email, or Facebook to verify their identity. Some participants were asked to submit an appropriate form of identification, but there was generally a lack of response or participants were unable to provide the required documentation. Because participants contacted us via Facebook private messaging to receive their coupon codes and their gift cards, we also verified their profile information against the identification documents provided. For those where suspicion was raised or if no identifying information was provided, Facebook accounts were often found to be recently created with no content or friends, confirming a fake profile. If fraudulent behavior was confirmed, the participant was sent a message informing them that they were ineligible to continue to participate in the study or to receive their incentives. If a participant acted as their own referral and completed the survey multiple times, only the first of the duplicated surveys was retained for statistical analysis.

Data Analysis

Statistical data analyses were performed using SPSS version 21 (IBM Corp, Armonk, NY, USA). Pearson chi-square (χ^2) test (Fisher’s exact test if applicable) was used to assess the difference in sociodemographic profile and drinking behavior between data collected by RDS and WebRDS techniques. All tests were 2-tailed using a significance level of 5%.

Demographic characteristics (gender, age, indigenous status, main language spoken at home, and socioeconomic status [SES] measured by postcode) and drinking behaviors of the sample recruited via RDS were compared to those sampled by WebRDS. Participants’ socioeconomic level was derived from population data from the 2011 Australian Socio-Economic Indexes for Areas (SEIFA) summary by postal area by Relative Socio-Economic Advantage and Disadvantage. The SEIFA provides population estimates by ranking them on a scale of advantage (high values) to disadvantage (low values) [29].

Cost Analysis

In calculating the total cost of traditional RDS, conservative estimates of an average of 3 hours per agency or location, with 2 study staff per visit receiving AU $30 per hour, were used. We consider this conservative because liaison with agencies was not incorporated. This liaison occurred for 6 weeks before data collection began and continued during data collection for another 6 weeks. The costs account for travel, estimating an average distance of 30 kilometers to and from each location by each study staff member, reimbursed at a rate of AU 64.5 cents per kilometer. We estimated an additional 30 minutes was spent by paid staff for administration purposes, namely sending 3 sets of follow-up referral messages to each participant recruited by RDS and WebRDS. We estimated this was approximately 300 hours of work, which corresponded to AU $9000 in staff costs. Further, the RDS costs did not include the time spent by study staff members to follow up agencies before visiting or afterwards to arrange additional visits. The expenses for WebRDS included the cost of the Facebook campaign (AU $430.44), with conservative estimates of one full-time equivalent staff member (38 hours per week) employed to monitor data for 8 weeks. Finally, incentives were not costed in this analysis for either method.

Results

Respondent-Driven Sampling Seed Recruitment

Using the described methods, a total of 148 organizations were initially contacted, of which 72 were sports clubs and 76 youth programs. Of these, 25 organizations agreed to participate and received presentations facilitated by research staff. Ten seeds were successfully recruited from 9 sports clubs and 11 seeds from 10 youth programs, all of whom completed the survey. In addition, 54 seeds were recruited across 20 community locations and via study promotion flyers or referrals by parents or teachers, with a total of 75 seeds recruited.

Web-Based Respondent-Driven Sampling Seed Recruitment

A total of 68 seeds were recruited (61 through Facebook and 7 from friends liking the study page or posts).

Advertising Campaign

To measure the reach of the Facebook campaigns, “impressions” (the number of times an advertisement was displayed to members of the target demographic), “reach” (the number of people who received impressions of an advertisement or page post), and “clicks” (the number of clicks an advertisement received) were recorded. The 3 advertisements, 6 promoted posts, and 11 status updates resulted in a total of 652,522 impressions, a reach of 88,280 adolescents, and 1426 (1.62% of possible accounts reached via the campaign) youth clicked the advertisement or post. Five of the 6 posts were targeted to youth aged 14 to 17 years Australia-wide with only the final post specific to the Perth region. The number of youth who took action in response to an advertisement or page post totaled 1412 (1.6%), which included actions such as page likes, conversations, and post comments. Figure 1 describes this recruitment process. The highest reach was achieved by a post that received 17,641 impressions resulting in 664 response actions. Generally, posts had lower impressions but a higher numbers of clicks, which is likely due to the larger group that could see the posts (Australia-wide) compared to the advertisements (Perth region).

Recruitment Results

A total of 96 surveys were excluded. These included surveys from respondents who reported the same contact details across multiple survey entries (n=47) and provided their parent’s contact details to do the survey multiple times (n=5). Participants who provided incorrect demographics were also excluded if they did not complete the survey within Perth and participants were unable to verify they lived in Perth (n=21), or respondents who were suspected to have reported a false age and did not provide a form of verification of identity on request (n=23).

A total of 143 seeds (75/143, 52.4% in RDS; 68/143, 47.5% in WebRDS) were recruited and completed the survey in both sampling methods, resulting in 869 valid participants recruited through the referral process (157/869, 18.1% in RDS; 712/869, 81.9% in WebRDS) (Figure 2). Table 1 presents the final sample of valid seeds that completed the survey. Seeds represented 14.13% (143/1012) of the overall sample recruited; 69.2% (99/143) of all seeds recruited made at least one referral and 46.25% (468/1012) of all participants recruited peers to the study.

The mean chain length in this study was 2.3 (SD 2.6), ranging from 0 to 12 waves per chain. When excluding seeds who did not recruit any participants (n=37), the mean chain length was 3.1 (SD 2.5). The majority of chains (n=37) consisted of one recruited wave followed by 2 waves (n=21) and 3 waves (n=12) per chain. In all, 25 chains recruited between 4 and 7 waves, whereas fewer chains (n=7) consisted of waves ranging from 8 to 12. The lengths of RDS and WebRDS chains did not vary significantly (*P*=.14).

Table 1. Characteristics of seeds who made a referral (n=99) and total recruited seeds (n=143) by recruitment method.

Age by recruitment method	Seeds, n (%)			
	Seeds who made referral		Total recruited seeds	
	Male	Female	Male	Female
RDS [a] (n=57 and n=75)				
14 years	7 (12)	5 (9)	4 (10)	2 (5)
15 years	8 (14)	6 (11)	8 (19)	3 (7)
16 years	6 (11)	9 (16)	5 (12)	8 (19)
17 years	6 (11)	10 (18)	8 (19)	4 (10)
WebRDS [b] (n=42 and n=68)				
14 years	9 (12)	7 (9)	6 (9)	3 (4)
15 years	9 (12)	11 (15)	11 (16)	8 (12)
16 years	8 (11)	11 (15)	6 (9)	12 (18)
17 years	8 (11)	12 (16)	10 (15)	12 (18)
All (n=99 and n=143)				
14 years	11 (11)	7 (7)	15 (10.5)	10 (7.0)
15 years	16 (16)	9 (9)	20 (14.0)	19 (13.3)
16 years	11 (11)	17 (17)	14 (9.8)	23 (16.1)
17 years	14 (13)	14 (14)	18 (12.6)	24 (16.7)

[a] Participants who (1) were recruited and completed survey face-to-face at sports clubs, youth programs, or community locations; (2) were recruited face-to-face and sent survey link and password following expression of interest to Facebook business page; and (3) expressed interest to Facebook study page or email contact to research staff as a result of seeing a study flyer or being made aware of the study through friends, teachers, or parents.

[b] Participants who expressed interest on Facebook study page as a result of viewing a study advertisement or a friend's interaction with the study business page, and who completed the survey on provision of the survey link and a password by the agency coordinator, whom we had liaised with.

There was an approximately 5-fold increase in rates of data collection with the transition from traditional RDS to WebRDS (see Figure 3). In the initial 35 days of traditional RDS, a mean of 2.2 (SD 5.0) surveys were completed per day compared to a mean of 8.9 (SD 10.2) surveys per day after WebRDS was launched. The peak daily rate of data collection, which can be attributed solely to completion of online surveys, was 41, which corresponded with the placement of a Facebook advertisement on a Friday. The peak combined rate occurred on the following Monday when 30 surveys were completed online and an additional 15 were recruited by face-to-face recruitment.

Cost Analysis Results

WebRDS allowed us to boost the number of surveys and speed of data collection by including Facebook posts and advertisements, which resulted in an immediate amplified response. Within RDS, differences were noted regarding recruitment at agencies compared to community locations, the former being more labor-intensive. This was due to the process of initially contacting the agency to establish interest, organizing and conducting the presentation, and then arranging a subsequent time for data collection if it was not feasible immediately after the presentation. In contrast, recruitment at community locations simply involved coordinating a location and time between study staff, conducting brief overviews of the study, and administering surveys with a seed and one or more of their referrals.

Figure 2. Recruitment process.

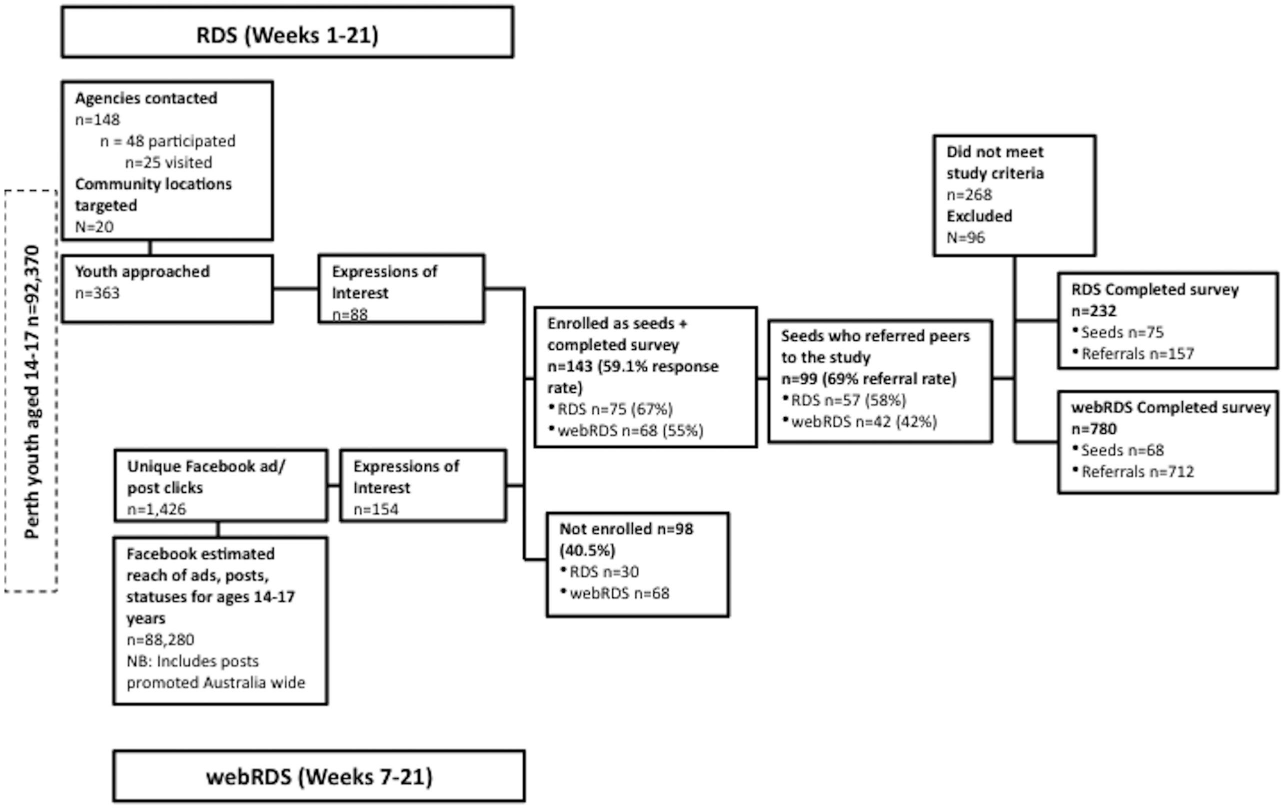


Figure 3. Differences in RDS and webRDS data collection pace.

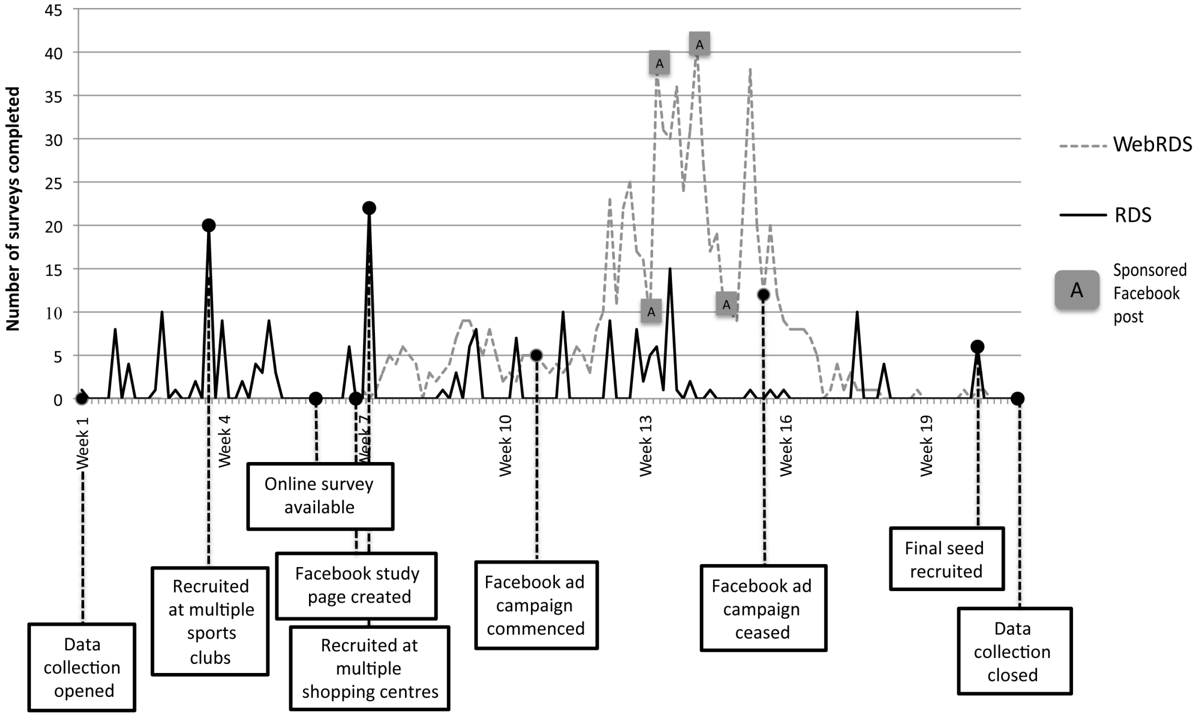


Table 2 provides a simple calculation of costs related to each survey method. The average expenses per RDS participant based on the calculations in Table 2 were AU $53.41 compared to AU $27.24 per WebRDS participant.

Table 2. Calculation to compare cost per survey of RDS with WebRDS

Cost per recruitment method	Formula	Calculation (AU$)	Total cost (AU$)
Total cost			
RDS	(Number of agencies and locations[a] visited*mean time spent at each location*number of study staff*cost to pay 1 study staff per hour)+number of agencies and locations[a] visited*number of study staff*(mean distance traveled to location*cost per km of travel)+number participants recruited by RDS*administration time per participant*cost to pay 1 study staff per hour	(45*3*2*30)+(45*2*30*.65)+(232*0.5*30)	13,335
WebRDS	Cost of Facebook campaign+(hours of work by 1 staff member*number of staff*cost to pay 1 study staff per hour)+number participants recruited by WebRDS*administration time per participant*cost to pay 1 study staff per hour	430.44+((38*8)*30)+780*0.5*30	21,350.44
Per participant			
RDS	Total cost of RDS/number of participants (seeds and referrals) recruited by RDS	12,390/232	53.41
WebRDS	Total cost of WebRDS/number of participants (seeds and referrals) recruited by WebRDS	21,250.44/780	27.24

[a] Number of agencies and locations visited were n=25 and n=20, respectively.

Participant Description and Drinking Behavior

A total of 1012 (n=232 in RDS; n=780 in WebRDS) valid surveys were included in the analysis. A similar proportion of gender and age groups were recruited in both samples (see Table 3). RDS recruited a significantly higher proportion of youth who identified as Aboriginal or Torres Strait Islander (5.6%, 13/232 RDS vs 0.8%, 6/780 WebRDS, *P*<.001), lower proportions of youth who spoke other languages than English as their main language at home (5.6%, 13/232 RDS vs 10.5%, 82/780 WebRDS, *P*=.03), and tended to recruit more adolescents living in areas of middle and lower SES (33.6%, 78/232 RDS vs 17.4%, 136/780 WebRDS, *P*<.001) compared to WebRDS. Significant differences in some alcohol use characteristics (past 7-day drinking and past 7-day risky drinking) were observed between the sampling methods, but these differences need to be treated with caution due to the low sample sizes.

Discussion

Principal Results

This study recruited a sample of 1012 adolescents in the community, of which 232 were recruited by RDS and 780 by WebRDS. In summary, a significantly larger proportion of Aboriginal or Torres Strait Islander (*P*<.001) participants who spoke English as their main language at home (*P*=.03) and of middle and lower SES (*P*<.001) was found in the RDS sample. The RDS sample was also found to have a higher occurrence of past 7-day drinking (*P*<.001) and past 7-day risky drinking (*P*=.004). No significant differences in gender, age, past month alcohol use, and lifetime alcohol use were observed between the RDS and WebRDS samples. This study revealed RDS and WebRDS used similar lengths of chains for recruiting participants; however, WebRDS had a faster rate of recruitment at a lower average cost per participant compared to RDS.

Table 3. Demographic characteristics and alcohol prevalence rates for RDS and WebRDS samples.

Demographic variables	Sample population (unweighted), n (%) N=1012[a] RDS sample n=232[b]	 WebRDS sample n=780[b]	*P*
Sex			.82
Male	136 (58.6)	465 (59.6)	
Female	96 (41.4)	315 (40.4)	
Age (years)			.10
14	49 (21.1)	133 (17.1)	
15	71 (30.6)	207 (26.5)	
16	58 (25.0)	256 (32.8)	
17	54 (23.3)	184 (23.6)	
Indigenous status			<.001
Aboriginal or Torres Strait Islander	13 (5.7)	6 (0.8)	
Not Aboriginal or Torres Strait Islander	215 (94.3)	770 (99.2)	
Main language spoken at home			.03
Australian	218 (94.4)	694 (89.4)	
Other	13 (5.6)	82 (10.6)	
Socioeconomic level (SEIFA deciles)			<.001
1-3	19 (8.4)	5 (0.7)	
4-7	59 (26.2)	131 (17.1)	
8-10	147 (65.3)	630 (82.2)	
Drinking behavior			
Never used alcohol	51 (22.2)	210 (26.9)	.17
Past month drinking[b]	75 (64.1)	156 (54.4)	.08
Past 7-day drinking	50 (21.6)	90 (11.6)	<.001
Past 7-day risky drinking[c]	30 (13.0)	52 (6.7)	.004

[a] Numbers may not add up to total (N=1012; RDS: n=232; WebRDS: n=780) due to missing data.

[b] “Past month drinking” only completed by participants who reported having ever drunk before.

[c] Risky drinking derived from “past 7-day drinking” variable using National Health and Medical Research Council guidelines of >4 drinks per day [30].

Comparison of Respondent-Driven Sampling to Web-Based Respondent-Driven Sampling

WebRDS represented a considerable advantage over the traditional RDS recruitment method as evidenced by the significant increase in referrals and reduction in cost per participant. It is possible that the slow recruitment rate of traditional RDS was due to the incentive not matching perceived effort and time required, or the intimidating aspects of contacting and meeting unfamiliar study staff. WebRDS reduced staff requirements and was more conducive to the online presence and preference for interaction via social media of the target group. In contrast to phone contact, interaction via Facebook appeared to be easier and possibly less invasive to study participants. Further, WebRDS enhanced anonymity and allowed participants to complete the survey in their own personal space and separate from peers. Most referrals were made in the first 2 weeks following survey completion, supporting the importance for rapid follow-up with participating youth in an appropriate manner to maximize chances of referrals being made.

In this study, we managed to recruit a large number of community-based participants within a short period after exploring various recruitment options. Significant human resources were required to conduct both the RDS and WebRDS recruitment. WebRDS yielded a considerably higher rate of completed surveys in a much shorter timeframe and at approximately half the cost per participant than traditional RDS. This is similar to findings of Wejnert and Heckathorn [24] suggesting that WebRDS has the potential to recruit large study samples up to 20 times faster than traditional RDS. Bauermeister and colleagues [21] also reported an expedited recruitment rate after offering participants the option to refer peers by email, text message, or social network post or message. However, there are differences in the methods implemented that limit direct comparison. The latter study recruited nearly 3500 participants in a 6-week period. This was facilitated by the use of automatically generated coupon codes and referral emails allowing a very quick and efficient response to participants. In contrast, we conducted the process manually, which left some participants waiting up to 48 hours for a response. Further, Bauermeister et al [21] allowed payment for up to 5 referrals with multiple use of referral codes and sampled from various regions across the United States, not just one city.

WebRDS proved to be an efficient method of recruitment and data collection with significant advantages over the RDS in agency and community locations. Notwithstanding, and similar to Bauermeister et al [21], a number of issues arose using this method. WebRDS posed a higher risk for duplicate and falsified surveys during the online recruitment and a substantial amount of time was allocated to screening participants and surveys to verify the data validity. Immediacy of response to potential participants and recruits to mail out survey invitations, referral messages, and incentives was vital to maintain the recruitment momentum and daily data screening and participant follow-up was necessary to ensure the quality of the data. However, this was difficult to achieve during peak times of survey completion.

Other studies using Facebook advertisements and online surveys reported confirming eligibility only as part of their survey [31-33]. However, due to the challenges encountered with online recruitment and confirming a person's identity, more rigorous procedures were adopted in this study. Although there was a risk of deterring youth from taking part by requesting identification, we included this step to ensure validity of the data. This was particularly important because most of the participants who expressed interest were seeds, which could have determined the quality of data for an entire chain of participants.

Almost half of our sample (46.25%, 468/1012) recruited participants to the study. These results are comparable to other RDS studies and coherent with the geometry reported for RDS recruitment patterns [24]; namely, if each respondent is asked to recruit 3 peers, approximately one-third of participants will make a referral [24]. For example, Thompson et al [34] studied street-involved youth aged 14 to 24 years using RDS; of 156 participants who received referral coupons, 67 (42.9%) recruited at least one peer. Of the 468 respondents who made a referral, 99 represented seeds resulting in 9.8% of the total sample recruited. Other studies have reported varying results with the proportion of seeds in relation to the total sample recruited, ranging from 1.9% [35] to 59.7% [34], possibly owing to the frequently reported difficulties in recruiting active seeds [36,37]. The length of a chain indicates the success of RDS and it has been proposed that social connectedness exists between participants if at least one chain achieves to recruit 3 waves [34]. No differences were found between the number of waves per chain between RDS and WebRDS (*P*=.14) in this study, with the mean length of chains being 2.3 (SD 2.6) waves and a maximum of 12 waves reached per chain. A large proportion of chains consisted of 1 to 2 waves only (n=58); however, we also managed to recruit 12 chains consisting of 3 waves and 32 chains, which achieved 4 to 12 waves, suggesting that the definition of connectedness between participants was met. There is limited information reported on the lengths of chains recruited by other RDS studies with young people; however, the findings suggest that our study recruited substantially more chains with longer waves [14,34,38]. For instance, Thompson et al [34] recruited a total of 17 chains of which only 3 achieved 3 to 4 waves and 2 chains reached 7 and 9 waves, respectively.

In this study, WebRDS and traditional RDS recruited participants with similar proportions of males and females (*P*=.10), of age groups from 14 to 17 years (*P*=.82), and of "never consume alcohol" (*P*=.17) and of "consumed in past month" (*P*=.08). However, traditional RDS recruited more indigenous participants (*P*<.001), fewer non-Australia ethnicities (*P*=.03), and more participants at a lower socioeconomic level (*P*<.001) compared to WebRDS. Traditional RDS participants were also more likely to have consumed alcohol in the past 7 days (*P*<.001) and consumed alcohol at risky levels in the past 7 days (*P*=.004), although these results may not be accurately compared due to the low numbers in these categories.

Research from the United States reported that although nearly all teenagers aged 12 to 17 years use the Internet, those who do not are more likely to be in households with lower income and less access to technology [39]. Higher SES predicts current Internet use and amount of Internet use, with teenagers of higher SES more likely to use the Internet for more time [40] and more likely to use it for social purposes [41] than lower SES teenagers. Congruent with these findings, the WebRDS sample in our research had a higher average SES. If people with higher SES used the Internet more often (assuming this follows with more Facebook use) than those of lower SES, they were more likely to see the Facebook advertisements, posts, and statuses and more likely to receive their online referral. The agencies and community locations in our RDS sample varied with respect to SES with methods in place to ensure distributive representation. This likely contributed to the higher proportion of lower SES youth in RDS than WebRDS, reflecting the increased diversity of this sample. In addition, indigenous Australians are more likely to be lower SES [42] than nonindigenous Australians, which may explain why RDS sampled a higher proportion of indigenous youth than WebRDS.

There are several explanations that could be posited for the differences in participants' main language spoken at home recruited by the different methods, whereby RDS recruited a smaller proportion of this group. Firstly, it could be because of the type of agencies and locations where recruitment occurred. Australian data show that children younger than 15 years born in non-English-speaking countries or whose parents were both born in non-English-speaking countries participate in organized sport at lower rates than their Australian-born counterparts [43]. Sports clubs accounted for 35% of the agencies visited in this study. In addition, the majority of sporting agencies were Australian Rules Football clubs, a sport primarily played in Australia, with less representation from more global sports such as cricket, gymnastics, and martial arts. Secondly, RDS recruitment was limited in agencies or locations that may have been more culturally diverse, such as at churches and specific cultural events. Finally, research shows that more children younger than age 15 years born in an English-speaking country other than Australia access the Internet than children born in Australia [43]. This is in contrast to what was observed by Bauermeister et al [21] that ethnic minorities use the Internet less than US-born individuals [39,40]. The combination of these factors led to non-English first language speakers being less likely to be sampled in RDS and more likely to be sampled in WebRDS.

Our research targeted specific groups through traditional RDS. There is potential that purposive seed sampling in WebRDS could be used to achieve a similar sample distribution as RDS. This would require more precise eligibility screening of potential seeds to ensure representation of different demographics. This could be achieved through Facebook when the person expresses interest (eg, screening on their age, postcode, and ethnicity). Doing this would come with its own set of challenges and may result in a smaller sample size depending on available resources, which would also limit representativeness. Future research on using the Internet, in particular Facebook, as a recruitment tool needs to focus on the best way of doing so to obtain high validity and reliability with careful consideration of the target population.

Facebook Recruitment

Facebook was used in this study to recruit participants to supplement conventional data collection methods. Using Facebook to enhance our recruitment resulted in fast response rates and a wide reach, presumably due to being a more acceptable form of communication among youth [44]. This has important implications because previous research has noted barriers to recruiting adolescents for research studies [33]. However, a recent Australian study recruiting young women aged 18 to 23 years via a range of different methods reported Facebook achieved the greatest success, recruiting a cohort of young women with similar characteristics to those of Australian women in terms of age, area of residence, and relationship status [45].

Although young people's access to the Internet is high in Australia [43], sampling from Facebook could have introduced some biases because the population was limited to adolescents who have access to the Internet, have a Facebook account, provided and matched the demographics (eg, age and location) targeted by our advertisements, and were actively logged into their account while the advertisements were screened. There may also be issues when targeting those with registered "interests" which may be more effective in attracting those interested in the topic [46,47]. Yet, this method could also prove useful when aiming to study subpopulations with specific attitudes and behaviors as was the case in our study.

Interestingly, our advertising campaign achieved a higher rate (1.6%) of youth who took action in response to an advertisement or promoted post in proportion to reach compared to other research using Facebook for recruitment. For example, Kapp and colleagues [31] and Ramo and colleagues [46] reported 0.075% and 0.7% of potential accounts reached via their campaigns resulted in clicks, respectively. These variations may be due to different incentives, target groups and cultures, Internet access behaviors, and that the use of posts in our study may have reached more users than advertisements only. Due to the peer referral process, we were unable to determine how most participants found out about the study because we were concerned that it would have further increased respondent burden and deterred participants. Thus, in most cases it was not possible to elucidate whether the advertisements/posts participants saw were screened by our campaign or were viral posts from Facebook friends.

Limitations

Overall, there were difficulties in motivating adolescents to take part in the study. There appeared to be a greater interest in the study among males, which was reflected in their rate of participation. It is possible that the form of incentive was more appealing to males, thus more neutral incentives should be considered in future studies in which gender equality is desired. As part of the overall project objectives, during the RDS, we specifically sampled youth from community programs catering for at-risk youth. In contrast, the lack of specificity using WebRDS may have created over- or undersampling of certain minority groups.

In addition, although there were no significant differences in the length of RDS and WebRDS chains, the chains created by youth recruited in person who referred their peers "on the spot" may have represented different types of connections than those who invited peers via Facebook where any of their Facebook friends could have clicked on the survey link. Hence, the nature of connections between RDS and WebRDS participants may be inherently different in yet unexplored aspects. Due to the difficulties in accurately determining fraudulent activity in WebRDS, it is possible that not all duplicate or falsified surveys were detected. Future research should also note that surveys should be considered on a case-by-case basis. Nevertheless, our findings related to RDS and WebRDS recruitment contribute to the literature and provide a reference for others intending to conduct similar studies.

Conclusions

There is a need to constantly improve the quality of Web-based surveys [48]. This is one of the first papers to describe the processes undertaken to gain 2 samples using both traditional RDS and WebRDS. Overall, Facebook was the most successful recruitment source for adolescents to complete an online survey compared to face-to-face recruitment and other forms of online recruitment and referral. A factor that likely contributed to this is the increasing preference of social networking sites for communication purposes among youth, which reduce the barriers to participation than more traditional recruitment methods. Although we were successful in using Facebook as a recruitment strategy, it is still a novel method and more research is necessary to overcome associated challenges and minimize biases. WebRDS requires continual monitoring and cleaning of data to screen suspicious participants. Such monitoring and the need for quick responses can be challenging, particularly if all communications and generation of referral codes is done manually. However, WebRDS allowed for a faster rate of recruitment at a lower average cost per participant than traditional RDS. WebRDS increased the ease of informing the target population about the study and is particularly useful for recruiting populations, which are traditionally difficult to access.

Measures need to be in place to ensure the demographics of WebRDS match traditional RDS, which could be done by purposive seed selection in both methods. Many other popular social networking sites exist that other population groups may access in preference to Facebook and the key is to ask the desired target group their preferences. The experience of this study does not just promote Facebook as a recruitment tool, but is a cue to exploring social networking sites as a means of recruitment.

Acknowledgments

This study was supported by a Healthway (the Western Australian Health Promotion Foundation) research project grant (grant number: 20331). The funding body had no role in reviewing or approving the manuscript for publication. The authors thank Jonathan Hallett, Jude Comfort, and Satvinder Dhaliwal for their contributions to the design and implementation of the study. We also thank Laurissa Tieleman, Emma Adams, Jasmin Jau, Nina te Pas, Lucinda Watts, Yolexis Hernandez Aguilera, and Marc Zen for their help with recruiting agencies and data collection. The authors are grateful to the key health industry stakeholders for supporting this research: the Drug and Alcohol Office, Department of Education and Training, Department of Sport and Recreation, Western Australian Sports Federation, and Western Australia Police.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Application of the Checklist for Reporting Results of Internet E-Surveys (CHERRIES)[48] to the Youth Alcohol Norms Survey (YANS).

[PDF File (Adobe PDF File), 114KB - jmir_v17i12e285_app1.pdf]

References

1. Abdul-Quader AS, Heckathorn DD, McKnight C, Bramson H, Nemeth C, Sabin K, et al. Effectiveness of respondent-driven sampling for recruiting drug users in New York City: findings from a pilot study. J Urban Health 2006 May;83(3):459-476 [FREE Full text] [doi: 10.1007/s11524-006-9052-7] [Medline: 16739048]
2. Bengtsson L, Lu X, Nguyen QC, Camitz M, Hoang NL, Nguyen TA, et al. Implementation of web-based respondent-driven sampling among men who have sex with men in Vietnam. PLoS One 2012;7(11):e49417 [FREE Full text] [doi: 10.1371/journal.pone.0049417] [Medline: 23152902]
3. Evans-Campbell T, Lindhorst T, Huang B, Walters KL. Interpersonal violence in the lives of urban American Indian and Alaska Native women: implications for health, mental health, and help-seeking. Am J Public Health 2006 Aug;96(8):1416-1422. [doi: 10.2105/AJPH.2004.054213] [Medline: 16809604]
4. Gao X, Hamzah SH, Yiu CK, McGrath C, King NM. Dental fear and anxiety in children and adolescents: qualitative study using YouTube. J Med Internet Res 2013;15(2):e29 [FREE Full text] [doi: 10.2196/jmir.2290] [Medline: 23435094]
5. Magnani R, Sabin K, Saidel T, Heckathorn D. Review of sampling hard-to-reach and hidden populations for HIV surveillance. AIDS 2005 May;19 Suppl 2:S67-S72. [Medline: 15930843]
6. Prestage G, Brown G, De WJ, Bavinton B, Fairley C, Maycock B, et al. Understanding gay community subcultures: implications for HIV prevention. AIDS Behav 2015 Feb 26:2224-2233. [doi: 10.1007/s10461-015-1027-9] [Medline: 25711301]
7. Wang J, Carlson RG, Falck RS, Siegal HA, Rahman A, Li L. Respondent-driven sampling to recruit MDMA users: a methodological assessment. Drug Alcohol Depend 2005 May 9;78(2):147-157. [doi: 10.1016/j.drugalcdep.2004.10.011] [Medline: 15845318]
8. Wang J, Falck RS, Li L, Rahman A, Carlson RG. Respondent-driven sampling in the recruitment of illicit stimulant drug users in a rural setting: findings and technical issues. Addict Behav 2007 May;32(5):924-937. [doi: 10.1016/j.addbeh.2006.06.031] [Medline: 16901654]
9. Liu H, Liu H, Cai Y, Rhodes AG, Hong F. Money boys, HIV risks, and the associations between norms and safer sex: a respondent-driven sampling study in Shenzhen, China. AIDS Behav 2009 Aug;13(4):652-662. [doi: 10.1007/s10461-008-9475-0] [Medline: 18841459]
10. Lomba L, Apóstolo J, Mendes F. Drugs and alcohol consumption and sexual behaviours in night recreational settings in Portugal. Adicciones 2009;21(4):309-325 [FREE Full text] [Medline: 20011989]
11. Blay N, Calafat A, Juan M, Becoña E, Mantecón A, Ros M, et al. [Violence in nightlife environments and its relationship with the consumption of alcohol and drugs among young Spaniards]. Psicothema 2010 Aug;22(3):396-402. [Medline: 20667266]
12. Bjørkhaug I, Hatløy A. Utilization of respondent-driven sampling among a population of child workers in the diamond-mining sector of Sierra Leone. Glob Public Health 2009;4(1):96-109. [doi: 10.1080/17441690701464207] [Medline: 19153933]

13. Shahmanesh M, Wayal S, Cowan F, Mabey D, Copas A, Patel V. Suicidal behavior among female sex workers in Goa, India: the silent epidemic. Am J Public Health 2009 Jul;99(7):1239-1246. [doi: 10.2105/AJPH.2008.149930] [Medline: 19443819]
14. Gwadz MV, Cleland CM, Quiles R, Nish D, Welch J, Michaels LS, et al. CDC HIV testing guidelines and the rapid and conventional testing practices of homeless youth. AIDS Educ Prev 2010 Aug;22(4):312-327. [doi: 10.1521/aeap.2010.22.4.312] [Medline: 20707692]
15. Heckathorn DD. Respondent-driven sampling: a new approach to the study of hidden populations. Social Problems 1997 May;44(2):174-199. [doi: 10.2307/3096941]
16. Salganik MJ, Heckathorn DD. Sampling and estimation in hidden populations using respondent-driven sampling. Sociological Methodology 2004 Dec;34(1):193-240. [doi: 10.1111/j.0081-1750.2004.00152.x]
17. Kendall C, Kerr LR, Gondim RC, Werneck GL, Macena RH, Pontes MK, et al. An empirical comparison of respondent-driven sampling, time location sampling, and snowball sampling for behavioral surveillance in men who have sex with men, Fortaleza, Brazil. AIDS Behav 2008 Jul;12(4 Suppl):S97-S104. [doi: 10.1007/s10461-008-9390-4] [Medline: 18389357]
18. Ramirez-Valles J, Heckathorn DD, Vázquez R, Diaz RM, Campbell RT. From networks to populations: the development and application of respondent-driven sampling among IDUs and Latino gay men. AIDS Behav 2005 Dec;9(4):387-402. [doi: 10.1007/s10461-005-9012-3] [Medline: 16235135]
19. Stueve A, O'Donnell LN, Duran R, San DA, Blome J. Time-space sampling in minority communities: results with young Latino men who have sex with men. Am J Public Health 2001 Jun;91(6):922-926. [Medline: 11392935]
20. Berg S. Snowball sampling. In: Encyclopedia of Environmetrics. Chichester: J Wiley; 2002:8-532.
21. Bauermeister JA, Zimmerman MA, Johns MM, Glowacki P, Stoddard S, Volz E. Innovative recruitment using online networks: lessons learned from an online study of alcohol and other drug use utilizing a web-based, respondent-driven sampling (webRDS) strategy. J Stud Alcohol Drugs 2012 Sep;73(5):834-838 [FREE Full text] [Medline: 22846248]
22. Zablotska IB, Frankland A, Holt M, de WJ, Brown G, Maycock B, et al. Methodological challenges in collecting social and behavioural data regarding the HIV epidemic among gay and other men who have sex with men in Australia. PLoS One 2014;9(11):e113167 [FREE Full text] [doi: 10.1371/journal.pone.0113167] [Medline: 25409440]
23. Goel S, Salganik MJ. Assessing respondent-driven sampling. Proc Natl Acad Sci USA 2010 Apr 13;107(15):6743-6747 [FREE Full text] [doi: 10.1073/pnas.1000261107] [Medline: 20351258]
24. Wejnert C, Heckathorn DD. Web-based network sampling: efficiency and efficacy of respondent-driven sampling for online research. Sociol Method Res 2008 Aug 01;37(1):105-134. [doi: 10.1177/0049124108318333]
25. Zablotska I, Wit JD, Brown G, Maycock B, Fairley C, McKechnie M, et al. Protocol for a respondent-driven sampling study exploring the roles of peer norms in HIV-related practices of gay men. WJA 2013;03(04):320-326. [doi: 10.4236/wja.2013.34041]
26. Hildebrand J, Maycock B, Howat P, Burns S, Allsop S, Dhaliwal S, et al. Investigation of alcohol-related social norms among youth aged 14-17 years in Perth, Western Australia: protocol for a respondent-driven sampling study. BMJ Open 2013;3(10):e003870 [FREE Full text] [doi: 10.1136/bmjopen-2013-003870] [Medline: 24154519]
27. Hildebrand J, Maycock B, Comfort J, Burns S, Howat P, Allsop S. Ethical considerations in investigating youth alcohol norms and behaviours: a case for mature minor consent. Health Promot J Aust 2015:1-5. [doi: 10.1071/HE14101]
28. Hildebrand J, Maycock B, Burns S, Zhao Y, Allsop S, Howat P, et al. Design of an instrument to measure alcohol-related psychosocial influences in the development of norms among 13-year-old to 17-year-old adolescents. BMJ Open 2013;3(8):e003571 [FREE Full text] [doi: 10.1136/bmjopen-2013-003571] [Medline: 23959761]
29. Australian Bureau of Statistics. 2013 Mar 28. Census of population and housing: Socio-Economic Indexes For Areas (SEIFA), Australia, 2011, Postal Area (POA) (Cat no 2033.0.55.001) URL: http://www.abs.gov.au/ausstats/abs@.nsf/DetailsPage/2033.0.55.0012011?OpenDocument [accessed 2014-01-03] [WebCite Cache ID 6MLDCFFG6]
30. National Health and Medical Research Council. Australian Guidelines to Reduce Health Risks from Drinking Alcohol. Canberra: Commonwealth of Australia; 2009. URL: http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/ds10-alcohol.pdf [WebCite Cache ID 6MLDKA8kO]
31. Kapp JM, Peters C, Oliver DP. Research recruitment using Facebook advertising: big potential, big challenges. J Cancer Educ 2013 Mar;28(1):134-137. [doi: 10.1007/s13187-012-0443-z] [Medline: 23292877]
32. Williams MT, Proetto D, Casiano D, Franklin ME. Recruitment of a hidden population: African Americans with obsessive-compulsive disorder. Contemp Clin Trials 2012 Jan;33(1):67-75 [FREE Full text] [doi: 10.1016/j.cct.2011.09.001] [Medline: 21983626]
33. Fenner Y, Garland SM, Moore EE, Jayasinghe Y, Fletcher A, Tabrizi SN, et al. Web-based recruiting for health research using a social networking site: an exploratory study. J Med Internet Res 2012 Feb 01;14(1):e20 [FREE Full text] [doi: 10.2196/jmir.1978] [Medline: 22297093]
34. Thompson LH, Schellenberg J, Ormond M, Wylie JL. Social structural analysis of street-Iinvolved youth in Winnipeg, Canada. JoSS 2011;12(6):1-17 [FREE Full text]
35. Decker MR, Marshall BD, Emerson M, Kalamar A, Covarrubias L, Astone N, et al. Respondent-driven sampling for an adolescent health study in vulnerable urban settings: a multi-country study. J Adolesc Health 2014 Dec;55(6 Suppl):S6-S12. [doi: 10.1016/j.jadohealth.2014.07.021] [Medline: 25454005]

36. Wylie JL, Jolly AM. Understanding recruitment: outcomes associated with alternate methods for seed selection in respondent driven sampling. BMC Med Res Methodol 2013;13:93 [FREE Full text] [doi: 10.1186/1471-2288-13-93] [Medline: 23865487]
37. Hathaway AD, Hyshka E, Erickson PG, Asbridge M, Brochu S, Cousineau M, et al. Whither RDS? An investigation of Respondent Driven Sampling as a method of recruiting mainstream marijuana users. Harm Reduct J 2010 Jul;7(15):1-11 [FREE Full text] [doi: 10.1186/1477-7517-7-15] [Medline: 20618944]
38. Johnston L, Thurman T, Mock N, Nano L, Carcani V. Respondent-driven sampling: A new method for studying street children with findings from Albania. Vulnerable Child Youth Stud 2010 Apr 07;5(1):1-11. [doi: 10.1080/17450120903193923]
39. Lenhart A, Madden M, Hitlin P. Teens and Technology: Youth Are Leading the Transition to a Fully Wired and Mobile Nation. Washington, DC: Pew Internet & American Life Project; 2005 Jul 25. URL: http://www.pewinternet.org/~/media/Files/Reports/2005/PIP_Teens_Tech_July2005web.pdf.pdf [accessed 2013-12-10] [WebCite Cache ID 6LkikaH2E]
40. Borzekowski DL, Rickert VI. Adolescents, the Internet, and health: issues of access and content. J Appl Dev Psychol 2001 Jan;22(1):49-59. [doi: 10.1016/S0193-3973(00)00065-4]
41. Peter J, Valkenburg PM. Adolescents' internet use: testing the "disappearing digital divide" versus the "emerging digital differentiation" approach. Poetics 2006 Aug;34(4-5):293-305. [doi: 10.1016/j.poetic.2006.05.005]
42. Australian Bureau of Statistics. 2011 Feb 17. The health and welfare of Australia's Aboriginal and Torres Strait Islander peoples: demographic, social and economic characteristics: language, culture and socioeconomic outcomes, Australia (Cat no 4704.0) URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/lookup/4704.0Chapter230Oct+2010 [accessed 2014-01-03] [WebCite Cache ID 6MLDVJlPw]
43. Australian Bureau of Statistics. Australian Bureau of Statistics. 2012 Oct 31. Children's participation in cultural and leisure activities-Western Australia (Cat no 4901.0) URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4901.0Apr%202012?OpenDocument [accessed 2014-01-03] [WebCite Cache ID 6MLEAmWKt]
44. Yonker LM, Zan S, Scirica CV, Jethwani K, Kinane TB. "Friending" teens: systematic review of social media in adolescent and young adult health care. J Med Internet Res 2015 Jan 05;17(1):e4 [FREE Full text] [Medline: 25560751]
45. Loxton D, Powers J, Anderson AE, Townsend N, Harris ML, Tuckerman R, et al. Online and offline recruitment of young women for a longitudinal health survey: findings from the Australian Longitudinal Study on Women's Health 1989-95 cohort. J Med Internet Res 2015 May 04;17(5):e109 [FREE Full text] [doi: 10.2196/jmir.4261] [Medline: 25940876]
46. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012 Feb 23;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
47. Khazaal Y, van Singer M, Chatton A, Achab S, Zullino D, Rothen S, et al. Does self-selection affect samples' representativeness in online surveys? An investigation in online video game research. J Med Internet Res 2014 Jul 07;16(7):e164 [FREE Full text] [doi: 10.2196/jmir.2759] [Medline: 25001007]
48. 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] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]

Abbreviations

RDS: respondent-driven sampling
SEIFA: Socio-Economic Indexes for Areas
SES: socioeconomic status
WebRDS: Web-based respondent-driven sampling

Edited by G Eysenbach; submitted 18.06.15; peer-reviewed by Y Jayasinghe; comments to author 14.09.15; revised version received 09.10.15; accepted 16.10.15; published 24.12.15.

Please cite as:
Hildebrand J, Burns S, Zhao Y, Lobo R, Howat P, Allsop S, Maycock B
Potential and Challenges in Collecting Social and Behavioral Data on Adolescent Alcohol Norms: Comparing Respondent-Driven Sampling and Web-Based Respondent-Driven Sampling
J Med Internet Res 2015;17(12):e285
URL: http://www.jmir.org/2015/12/e285/
doi:10.2196/jmir.4762
PMID:26704736

Original Paper

Internet-Based Direct-to-Consumer Genetic Testing: A Systematic Review

Loredana Covolo[1], ScD; Sara Rubinelli[2], DPhil, PhD; Elisabetta Ceretti[1], ScD; Umberto Gelatti[1], MD

[1]Unit of Hygiene, Epidemiology and Public Health, Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, University of Brescia, Italy, Brescia, Italy

[2]Department of Health Sciences and Health Policy, University of Lucerne and Swiss Paraplegic Research, Lucerne/Nottwil, Switzerland

Corresponding Author:
Loredana Covolo, ScD
Unit of Hygiene, Epidemiology and Public Health
Department of Medical and Surgical Specialties, Radiological Sciences and Public Health
University of Brescia, Italy
Viale Europa 11
Viale Europa 11
Brescia, 25127
Italy
Phone: 39 030 3717697
Fax: 39 030 3717688
Email: loredana.covolo@unibs.it

Abstract

Background: Direct-to-consumer genetic tests (DTC-GT) are easily purchased through the Internet, independent of a physician referral or approval for testing, allowing the retrieval of genetic information outside the clinical context. There is a broad debate about the testing validity, their impact on individuals, and what people know and perceive about them.

Objective: The aim of this review was to collect evidence on DTC-GT from a comprehensive perspective that unravels the complexity of the phenomenon.

Methods: A systematic search was carried out through PubMed, Web of Knowledge, and Embase, in addition to Google Scholar according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist with the key term "Direct-to-consumer genetic test."

Results: In the final sample, 118 articles were identified. Articles were summarized in five categories according to their focus on (1) knowledge of, attitude toward use of, and perception of DTC-GT (n=37), (2) the impact of genetic risk information on users (n=37), (3) the opinion of health professionals (n=20), (4) the content of websites selling DTC-GT (n=16), and (5) the scientific evidence and clinical utility of the tests (n=14). Most of the articles analyzed the attitude, knowledge, and perception of DTC-GT, highlighting an interest in using DTC-GT, along with the need for a health care professional to help interpret the results. The articles investigating the content analysis of the websites selling these tests are in agreement that the information provided by the companies about genetic testing is not completely comprehensive for the consumer. Given that risk information can modify consumers' health behavior, there are surprisingly few studies carried out on actual consumers and they do not confirm the overall concerns on the possible impact of DTC-GT. Data from studies that investigate the quality of the tests offered confirm that they are not informative, have little predictive power, and do not measure genetic risk appropriately.

Conclusions: The impact of DTC-GT on consumers' health perceptions and behaviors is an emerging concern. However, negative effects on consumers or health benefits have yet to be observed. Nevertheless, since the online market of DTC-GT is expected to grow, it is important to remain aware of a possible impact.

(J Med Internet Res 2015;17(12):e279) doi:10.2196/jmir.4378

KEYWORDS

genetic testing; direct-to-consumer; Internet; online market; systematic review

Introduction

“There’s no gene for fate.” This is a quote from the movie “Gattaca,” a 1997 American science fiction film set in a future when one’s life is determined by genetic engineering rather than education or experience [1]. This theme expressed concern about the negative effects of a genetic determinism foreseen in a distant future. However, only a few years later, advertisements such as “Your future health is in your genes,” [2] “Your DNA, your personal health,” [3] or “Diet and exercise matched to your genes” [4] started to appear on websites of commercial companies offering direct-to-consumer genetic testing (DTC-GT). There are even companies offering tests to find genetic compatibility with a partner, presented as a key to successful and long-lasting romantic relationships [5]. One can imagine this scenario triggering a genetic determinism in potential consumers, mainly because there is no involvement from health professionals. The paradox is that, despite the fact that predictive genetic tests are already on the market, the majority of such tests lack scientific evidence and a proven clinical utility [6,7].

Over the past decade, the phenomenon of DTC-GT has generated a huge debate among physicians, bioethicists, and government bodies [8-12], and many recommendations are available [13-15]. In November 2013, the US Food and Drug Administration (FDA) ordered 23andMe, a provider of DTC genomic services, to stop marketing health-related genetic tests due to the risk that false results could cause consumers to undergo unnecessary health procedures [16]. However, there are currently other online companies offering this kind of service [2-4,17].

The current evidence on the risks of DTC-GT is uncertain. To our knowledge, there are three main systematic reviews on specific aspects related to DTC-GT [15,18,19]. These reviews, carried out by the same group of authors, separately explored the current position statements and recommendations on the use of DTC-GT [15], along with the views and experiences of consumers [18] and health professionals [19]. Analysis of documents produced by professional or public organizations [15] has caused great concern about potential harms for consumers who might undergo DTC-GT. Considering the difficulty in creating international standards that regulate the online market, the authors underlined the need to promote an agreement on a code of practice based on specific recommendations that include appropriate education for health professionals, as well as the guarantee of appropriate information to consumers. But there are mixed views on the actual risks of DTC-GT. With evidence that DTC-GT might actually increase the demand for consultation and related screening or diagnostic testing, some health professionals rated GT as clinically useful and a valuable opportunity for early screening [19].

There are two additional recent reviews on DTC-GT [20,21] that explore this topic in general and conclude that, from the consumer’s experience, there does not seem to be enough evidence to qualify the risks of these tests. Yet, these two reviews were not based on a systematic approach.

The objective of our review is to merge evidence on DTC-GT from a more comprehensive perspective than the studies mentioned above. In addition to identifying further literature on the value of DTC-GT from the point of view of consumers and health professionals, this review also considers the scientific evidence and clinical utility of this type of testing and the way DTC-GT is marketed from a health communications perspective. The analysis of these last two aspects are essential to offering a multifaceted framework for understanding the complexity of DTC-GT as a phenomenon and informing directions for future research and policy making in the field.

Methods

The systematic review was performed according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [22] (see Multimedia Appendix 1).

Information Sources

The literature search covered the period up to October 2014. The search was performed using electronic databases (PubMed, Web of Science, and Embase) and the Google search engine tool, Google Scholar. On Google Scholar, we investigated all the results obtained by the databases, but considered only the first 500 results because the number of relevant articles declined substantially after the first 300 results and because this search engine displays results by relevance using a link analysis system or algorithms [23].

Search Strategies

We used “Direct-to-consumer genetic test” as the key term for each database and for Google Scholar. We scanned the reference lists for relevant articles up to the second level, and we considered the “related articles” of relevant ones in the PubMed database or Google Scholar when the paper was not present in PubMed.

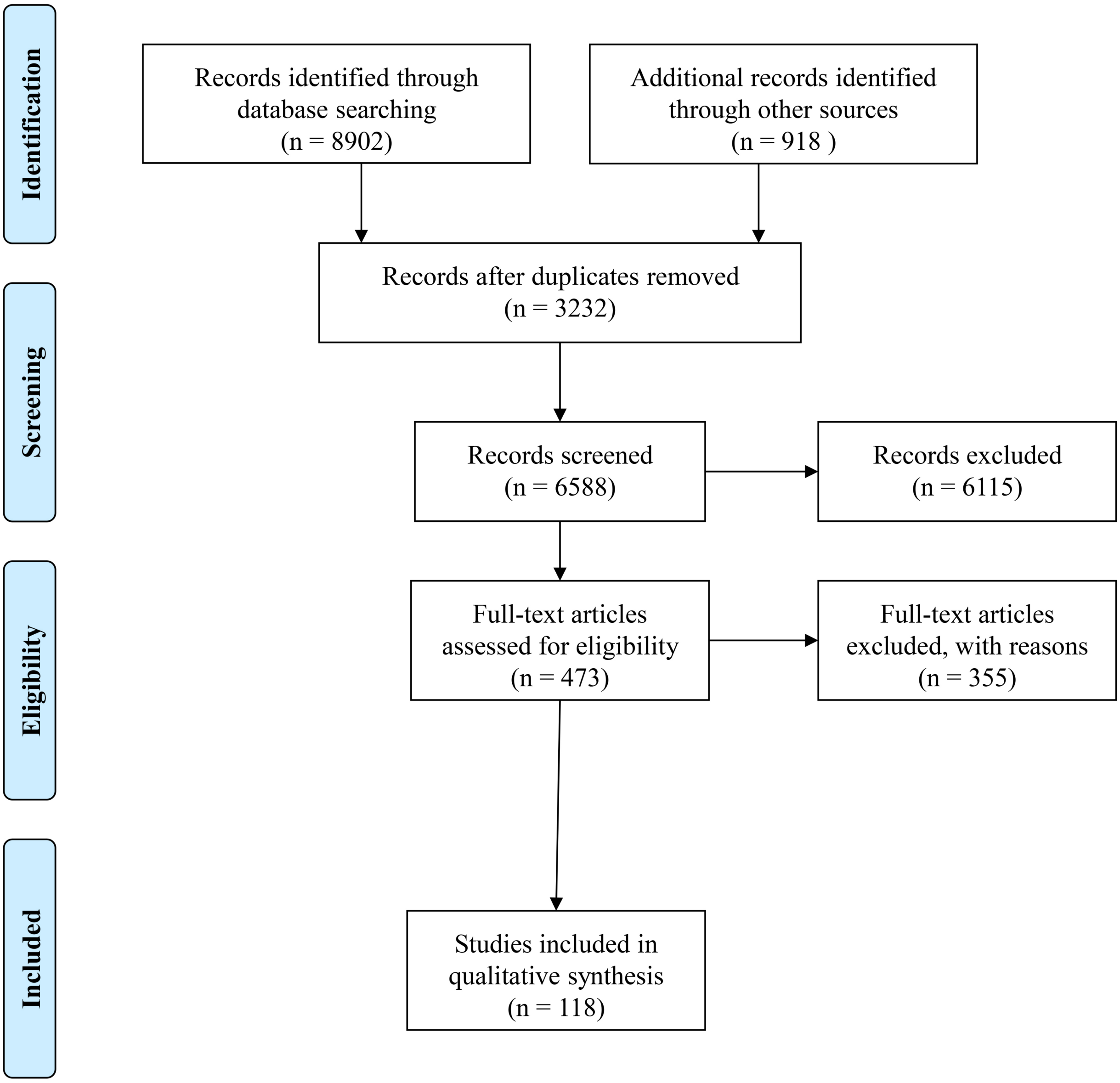
Study Selection and Eligibility Criteria

We included all articles relevant to the subject of the research where the key term was anywhere in the text of the paper, written in English, with the abstract and full text available. We included only scientific articles, excluding popular articles published in daily newspapers or weekly and monthly magazines. Papers included articles associated with health-related genetic tests available online and offered direct-to-consumer. We selected only the articles reporting original data, excluding those with speculative discussion about the problem or citing data from other studies (ie, editorials, letters, comments, articles about regulation issues, and reviews).

Two investigators read the papers (LC and UG) and independently assessed the potential relevance of all publications, identified during the database search, based on the information provided in the titles and abstracts. Disagreement was resolved by consensus.

After screening by titles and abstracts, the first author critically reviewed the full texts of the remaining articles and extracted the information required to perform the review. The methodological quality of each study was assessed by 2 authors (LC and EC) using the Kmet tool for evaluating quantitative and qualitative research [24]. A score of between 0 and 1 was assigned to each paper based on a series of questions related to the type of study. Case studies and descriptive reports (a total of 38 papers) were excluded from the evaluation. Disagreements were resolved through discussion among the authors until consensus was reached. As shown in Figure 1, we identified 118 articles that fit the inclusion criteria.

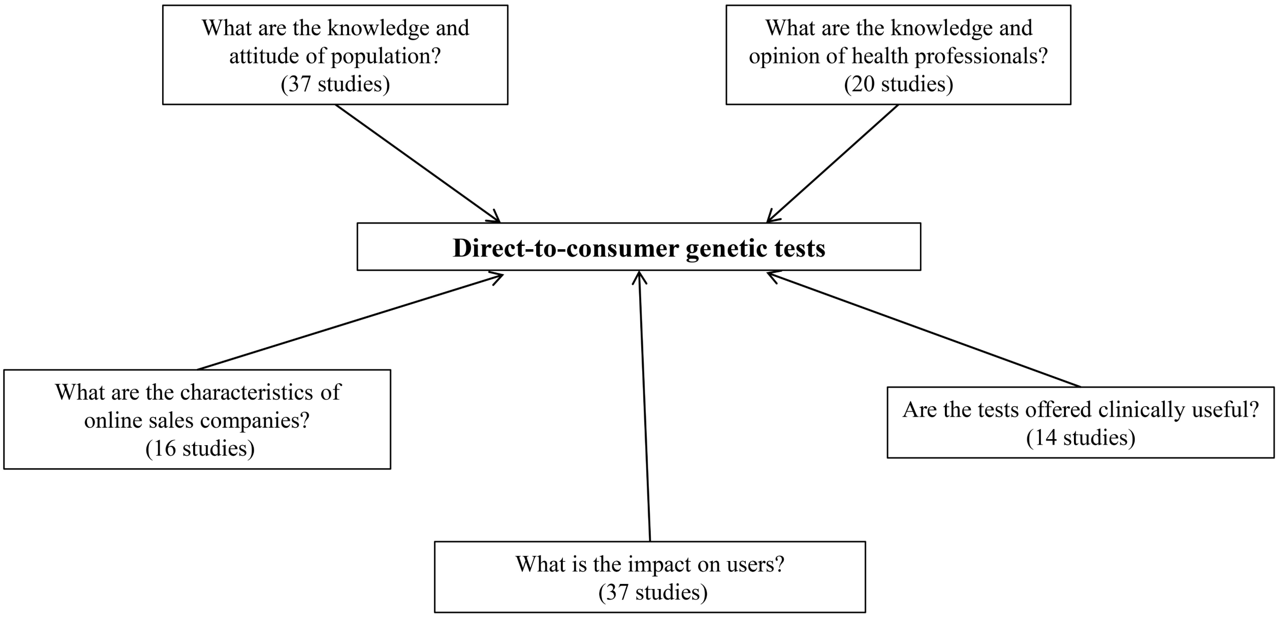
Figure 1. Flow diagram describing the study selection.



Results

Of the 473 relevant articles selected, we included 118 studies with original data (24.9%): 95 quantitative studies, 15 qualitative studies, and eight case studies. These articles have been divided in five categories as shown in Figure 2: (1) knowledge and attitude/perceptions to DTC-GT, (2) health professionals' opinions about DTC-GT, (3) characteristics of online companies selling GT, (4) DTC-GT's impact on users, and (5) evidence of clinical utility and validity. Some articles with original data covered more than one of these subjects and were consequently allocated to more than one group. For studies investigating DTC-GT's impact on users, we included studies investigating both hypothetical situations (n=20) and actual situations (n=17).

Figure 2. Research categories.



Research Categories

Public's Knowledge of and Attitude Toward Direct-to-Consumer Genetic Testing

We selected a total of 37 articles, of which four are qualitative studies, investigating the public's knowledge of and interest in DTC-GT [25-28] (see Table 1 in Multimedia Appendix 2). Of the 37 studies, 25 (68%) were carried out in the United States [25-50] with 41% (15/37) registering more than 1000 subjects [30-33,37,40,42,43,46,51-56]. The response rate was more than 60% in over half of the studies (16/30, 53%) where such data were available [30,32,33,35,36,38,40,45,47,50,52,56-60]. In general, the age range of participants was quite wide with a mean age ranging from about 25-50 years. The education level was high (college degree or more) for the majority of participants in all studies.

Overall, the level of awareness of DTC-GT was low, ranging from 8% [33] to around 50% [37,43,45,57]. In two studies with a large number of subjects (more than 4000) and a response rate of more than 60%, there was a very low level of knowledge of DTC-GTs, specifically 13% [59] and 14% [32].

A total of 78% of subjects recruited in the study by Gollust et al [34] were aware of personal genomics, but only 15% visited a DTC-GT website. A large portion of women heard about GTs (73%) in the study by Perez et al [47]. However, the sample was small (84 women) and characterized by women at high risk for breast cancer who may have been more aware of this subject.

In 4 studies [25,30,35,54], the participants expressed great interest in GTs. In fact, 82% of subjects recruited at the Scripps Transitional Science Institute reported that they would want to know their disease risk [30]. However, as highlighted by the authors, the sample was not representative of the general public because it was largely made up of Institute employees and a number of technology and biotechnology company employees. A large number of women indicated definite interest in GT (77%) in the study by Graves et al [35], but they were women at moderate to high risk for breast cancer. Similarly, the interest in having GT to determine susceptibility to major depression was higher in participants affected by a depressive disease than those unaffected (71% vs 64%), although not statistically significant [25]. In another seven studies [27,42,43,55,57,59,60], a moderate interest in GT (from 50-60%) was found. It should be noted that 48% of respondents in the Cherkas et al [59] study were interested in GT if the test was free of charge. Similarly, 37% of a sample of Canadian adults stated that they would pay nothing for GTs even if related to a manageable condition [53]. Only 5% were potentially interested at the current price (£250). In several studies, fewer than 40% of participants expressed favorable attitudes to GT [36,37,47,48,52,56,58].

The interest in GT seemed to increase only when the information received was positive [49], when people felt they would regret not taking the test [54] or, in the case of parents, when they could learn about their child's decreased risks [50]. Survey respondents who perceived greater threat from disease had significantly greater behavioral intentions to talk to their doctor and search for more information about the test, even if it did not affect their plans to take the test [61]. Additionally, when people were informed about the risks of DTC testing, they became less interested in getting GT [36]. At the same time, it was found that conscientiousness about the risk of GT, and not neuroticism, led people to seek online information about DTC-GT [46]. Web-based genomic information presented using evidence-based communications made patients more favorable to this type of testing [38].

The importance of having information about GT was also supported by the need to refer to a physician to interpret the test results [34,41,43,56]. Nearly half (46%) of women recruited in the Perez et al study [47] strongly agreed that it is more appropriate for companies to target doctors to identify women who may be at risk for carrying the breast cancer gene than target all women through different types of media. Respondents of a large Australian survey [51] were not comfortable with companies offering DTC-GT and were unlikely to order the test because it was perceived to be less regulated and accurate compared with a test provided by a conventional medical practitioner. Concerns about poor regulation of DTC-GT companies and violation of privacy emerged from media coverage of DTC-GT [26].

Among the reasons for favorable attitudes to GT were curiosity [34,43,59] and interest in monitoring and improving health [28,34,57,59]. University students in Switzerland reported the contribution to scientific research as their main reason for undergoing testing [55]. The availability of treatment was a factor that motivated the respondents of a Canadian sample population (61%), whereas curiosity had only a modest impact on willingness to pay for GT. Younger respondents were more likely to cite curiosity as a reason for testing [53].

Impact of Direct-to-Consumer Genetic Testing on Users

We retrieved a total of 29 articles dealing with the consequence of undergoing a DTC-GT, of which nine were qualitative studies [62-70] (see Table 2 in Multimedia Appendix 2). A qualitative analysis was also included in the study of Vayena et al [71]. Furthermore, there were eight case studies [72-79]. As much as 68% (25/37) of the studies, including case studies, were carried out in the United States [32,43,63-66,70,73,75-79, 81-93]. Excluding case studies, nine studies investigated the DTC-GT experience with people who actually purchased the test [32,43,62,67-69,90,91,94]. In the other studies, participants were investigated only as potential consumers. Overall, the main goal was to evaluate psychological reactions, behavioral effects, and perception risk.

In 41% of studies (12/29) [32,43,62-68,70,71,90], the sample size was very low (fewer than 100 subjects). There are some exceptions: Kaufman et al [91] recruited 1048 subjects (but with a response rate of only 33%) and Su et al [69]. Also, six studies [81-83,85,86,93] referred to the same large sample coming from the longitudinal cohort study of 3639 adults recruited from Scripps Health employees, employee family members, and Scripps Health patients who purchased the GT at a discounted rate [30].

Regarding the impact of GT results on health behavior, a large proportion of participants expressed the intention to modify lifestyle (eg, diet, exercise), both among actual customers [43] and hypothetical ones [55,63,80], and a modest change in health behavior was observed, particularly among people who purchased a DTC-GT [67,91,94]. In the study by Francke et al [90], 11 out of 16 women received information about being positive for the breast cancer type 1 susceptibility gene mutation (BRCA) from the DTC company. Although the number is modest, three of them had risk-reducing oophorectomy and four planned to. One had a mastectomy and three planned to. Five declared they went to have breast exams and breast imaging after getting their results.

No impact on potential user behavior was evidenced at the 3-month follow-up [93] or after a moderately longer period of observation (1 year) [65,82,92]. No influence on DNA-based dietary advice in personalized nutrition perception was observed in a randomized control trial after 1 year of follow-up [95].

Slightly more than half of the people who used a DTC-GT discussed their results with a physician [43]. Similarly, 60% of 23andMe customers, who showed as mutation-positive, reported sharing their results with their physician. Only 26% of the mutation-negative customers shared their information with their physician [90]. Increased physician utilization was found among people who underwent DTC pharmacogenetic tests [83].

In general, the number of people who reported sharing test results with a physician is quite low (<30%), both among actual users [32,67,91] and experimental ones [66,70,71,86], even though most participants stated that they would (or might) disclose to physicians when asked in the study by Wasson et al [66]. After the 1-year follow-up, no changes were found in the overall use of health care by those receiving personalized GT results compared to those who were not tested [92].

Generally, the proportion of people worrying about their tests results was also quite low. Fewer than 30% of DTC-GT customers declared a change in health anxiety [94] or felt anxious even if mutation-positive [90]. Furthermore, there was no significant difference between baseline and follow-up anxiety symptoms at 3 months [81,89] and 1 year [82,83,87] after receiving test results. Bloss et al [81,82] found that greater perceived seriousness and diminished perceived control over a disease were associated with test-related distress and higher, but not clinically significant, levels of anxiety [93]. In addition, people who shared their test results with health care providers were significantly less worried about being tested compared to non-sharers (45% vs 53%, *P*=.01), but only a small percentage (around 10%) were worried about learning of disease risk in both groups. Nevertheless, great value was attributed to risk information in 78% of sharers and 69% of non-sharers (*P*<.01) [86].

Several studies showed no concern by all or a majority of participants [55,63,65,71]. Concerns seemed to relate to the type of results. For example, a significant increase in negative effect was shown among individuals who learned that they were susceptible to alcoholism [84]. However, the intentions for alcohol consumption in the near future were not affected. In a study by Gordon et al [70], 88% of participants reported feeling reassured by these test results: indeed, they were encouraged by learning of their negative test results and their low-risk factor. Similar conclusions were found by Harris et al [62] who analyzed stories told by DTC-GT users. These participants even felt a sense of indifference toward the test results. On the contrary, almost half of people who knew about a cancer risk and 81% of people who learned about myocardial infarction risk through a DTC-GT were worried about these diseases. After 1 year, there were no differences in being worried compared to people who had not been tested [88].

People who did not interpret test results as deterministic on health outcomes, or declared they understood the results, were generally not worried about them [63,70,87,89]. People who were tested for four conditions perceived a higher risk than those who were not tested; a large portion of them even expressed concern about their disease risk. This difference was, however, not significant 1 year after receiving test results [88]. Only 10% of people interviewed in the study by Vayena [71] reported a serious impact on their health perception, 55% stated some impact, and 35% affirmed no impact, while less than half of these respondents reported having no concerns at all about DTC testing. It should be noted that their primary reported concerns regarded privacy issues.

In another context, it was interesting to find that many people signed a petition to support unrestricted access to DTC-GT stating also the health care professional and government should not be placed as intermediaries when purchasing DTC-GT [69]. The perception of having understood the test results was the main reason for not utilizing the counseling service [82]. However, the evidence showed that the delivery of personal genomic risk through a trained health professional resulted in significantly higher comprehension compared to online delivery [87].

An incorrect interpretation risk was found among people who underwent DTC-GT [91], and the majority of DTC consumers interviewed by McGuire et al [43] considered information obtained from DTC-GT to be a diagnosis of a medical condition. This evidence contrasts with other studies that revealed that many people were aware of the low predictive value of DTC-GT [67] or the fact that they report an average risk of disease [94]. The main reason for purchasing the GTs related to health, as well as a general curiosity about genetic make-up [43,62].

Case Studies

We retrieved seven studies reporting on patients who purchased a DTC-GT and one analyzing information from two reports from a DTC-GT company [73]. Except for two case studies in which DTC-GT were considered useful [77,78], all other case studies underlined the importance of correctly understanding and interpreting the results in order to avoid adverse psychological consequences [73,74,76], unnecessary preventive measures [79], or the possibility of giving the genetic profile a deterministic role [72]. This is particularly important when people learn about their susceptibility to Alzheimer's disease for which proven preventive strategies are still lacking. On the basis of 2 subjects who tested positive for Alzheimer's, the need to improve strategies for informed decision making was discussed. For example, DTC-GT could provide a more detailed consent form and promote a mandatory pre-test conversation with a genetic counselor. This is highly relevant for DTC-GT as the health information provided by DTC companies seems to be influenced by commercial loyalties and can therefore be potentially misleading [74].

Learning about a genetic predisposition to curable diseases may be beneficial, as was the case for a woman who learned from her DTC-GT that she was at high risk to develop breast cancer. She felt empowered by prevention, although it was genetic counseling that contributed to her facing and limiting her initial anxiety [75]. The support of a health professional is also crucial when considering the risk of misinterpreting the test results [73]. Another study reported the case of a 52-year-old man whose test results implied that his obesity was genetically predetermined and inevitable, but after appropriate lifestyle modification he lost 32 kg, indicating the importance of environmental factors [72]. In this context, Corpas [74] highlighted the need to have psychological support in sharing results with family, an aspect that is not emphasized in the DTC-GT process.

Health Professionals' Point of View on Direct-to-Consumer Genetic Testing

A total of 20 articles explored health professionals' opinions of DTC-GT as reported in Table 3 in Multimedia Appendix 2. Two of these are qualitative studies [96,97]. Out of 20 studies, 14 were mainly conducted in the United States (70%) [32,41,97-108]. Half of the surveys were implemented online [32,41,98-101,107-110].

Four studies recruited more than 1000 subjects [32,102,108,111], and five studies recruited more than 300 subjects [52,98,101,105,107]. Almost half of the studies (44%) had a response rate of around 40% [100-102,104,109,110,112] and other 6 studies (38%) had a response rate less than 20% [41,99, 103,105-107,111].

The awareness of DTC-GT among physicians was high (around 90%) in three studies [41,100,110]. In other surveys of physicians not specializing in genetics, fewer than 55% of respondents were aware of DTC-GT [102,105,111,112]. In general, the percentage of physicians who have discussed GT results with a patient or have ordered a GT for a patient was quite low. Sixteen percent of physicians recruited by Bernhardt et al [98] ordered one test a week or more, and only 7% reported having seen a DTC genetic risk assessment report. In five surveys, fewer than 19% of physicians, both specializing in genetics and not, reported having patients request genetic consultations [32,99,105,107,108]. Forty-four percent of clinical geneticists from different European countries had been contacted by at least one patient regarding DTC-GT services after purchasing the test [110], and 46% of genetic counselors from the United States had worked with patients who had initiated a discussion of DTC-GTs: however, only 15% of the latter had suggested them to their patients [101]. Forty-two percent of primary care physicians enrolled in an online survey in the United States had ordered a GT for a patient, and one third had ordered them for themselves [100]. Only 0.5% of general practitioners and 1% of clinical geneticists from Japan ordered DTC-GT [111]. A large percentage of neurologists (74%) and the 14% of psychiatrists recruited from the American Medical Association ordered a GT for a patient [107].

Another interesting topic is how confident physicians are in interpreting GT results. In the study by Bernhardt et al [98], 16% of primary care physicians declared themselves to be "very confident," along with 15% of family physicians in the study by Powell et al [106] and only 7% of physicians specialized in genetics from the study by Brett et al [109]. In a study on nutrigenomics, health professionals reported a lack of competency to provide information on nutritional genomics. Inability to support a patient in managing genetic risk information also emerged from interviews with 18 clinicians providing genomic risk assessment services to their patients [97]. The study by Salm et al [107] reported the need to have more training in interpreting GT results; although in the context of predictive genomic testing, the United States has promising training programs for genetic counselors [103].

Health professionals' opinions on the clinical utility of DTC-GT were contrasting, and the percentage of those who were in favor of these services was different among the studies. Giovanni et al [99] found that 52% of health care providers described the genetic test as clinically useful. The majority of respondents (86%) mentioned usefulness in the context of breast cancer susceptibility, in agreement with findings from Mainous et al [102]. The latter also found that 30% of participants perceived GT utility in detecting Alzheimer's disease and 25% of participants for heart disease or diabetes. Half of genetic counselors recruited by Hock et al [101] said that GT should be limited to a clinical setting, 23% of the sample was neutral, and 27% disagreed. Furthermore, 56% of the sample considered a DTC-GT acceptable only with the provision of genetic counseling, 31% were neutral, and 13% disagreed. In the study by Bernhardt et al [98], physicians thought that genetic tests would be helpful in managing patients; in particular, 70% felt it would be useful with pharmacogenomics and 40% with disease risk assessment. However, only one third of physicians in both cases would order such testing for a patient. About 47% stated that genetic testing would be helpful for patients, motivating them to adopt healthy behaviors. Also, the clinicians interviewed in a recent study [97] were enthusiastic about the potential of GTs to enhance the personalized, preventive, and wellness orientations of their clinical practices.

In a study conducted in Australia [109], the majority of genetic health professionals did not consider DTC-GT useful for individuals who want anonymous testing (54%), are driven by curiosity (54%), or are geographically isolated (60%). Forty-three percent of physicians in the Powell et al [105] study considered DTC-GT clinically useful. In a Greek study by Mai et al [52], only 13% of medical practitioners were in favor of DTC-GT. Similarly, 86% of clinical geneticists recruited from 28 European countries [110] considered it unacceptable to provide a predictive test without face-to-face medical supervision, and all respondents expressed the unacceptability of offering DTC-GT for conditions neither treatable nor preventable.

Ohata et al [111] carried out a survey on 1145 general practitioners and 294 clinical geneticists in Japan. Convenience scored highest in both groups as the reason behind users' ordering DTC-GT, and general practitioners rated the benefits of DTC-GT higher than clinical geneticists (score 2.54 vs 1.96 on a scale 1-4, 1=disagree). Among the risks, the concern for understanding results scored highest in both groups (score >3). Furthermore, reliability of results and provision of information/counseling were a source of concern greater in clinical geneticists than general practitioners (score 3.13 and 3.78 vs 2.77 and 3.48 respectively).

In a study conducted in New Zealand [112], general practitioners who had not received training in genetics agreed that convenience was a benefit, more than those with training (72% vs 38%, $P<.005$). At the same time, misinterpreting results and inadequate delivery of information were perceived to be the greatest risks associated with DTC-GT by the majority of respondents (around 90%). In general, only 19% agreed that DTC-GT provides a useful service in the delivery of health care, and 26% agreed that results encourage patients to take responsibility for their health. Clinical validity of the test (25%) and counseling (20%) were the most selected aspects regarding advertising regulation of DTC-GT.

In another study [108] dealing with GT in children, genetic counselors appeared less prone to GT compared with non-genetic physicians.

There was one study exploring the knowledge of and attitude toward personal genomics on a small group of medical students enrolled in a human genetic course [104]. The percentage of students who thought that genotyping information would be useful to physicians and consumers decreased after the course (32% post-course vs 63% pre-course and 52% vs 84% respectively). The majority of students, both before and after the course, expressed concerns about reliability and utility of results. They agreed that tests needed interpretation (around 70%) and DTC companies had to provide genetic counseling (71% pre- and 80% post-course).

More than 80% of physicians recruited by Powell et al [105] expressed concerns about possible misinterpretation of test results and increased anxiety in patients. Almost half of physicians (neurobiologists and psychiatrists) surveyed by Salm et al [107] thought that GT could cause psychological harm to their patients and they could be exposed to possible insurance discrimination. This was further confirmed in the study by Bernhardt et al [98].

Uncertainty about clinical utility concerned the majority of primary care physicians (around 60%) in the study of Haga et al [100], with a recommendation for health care professionals to act as intermediaries also when discussing DTC nutrigenomic tests [96].

Content of Websites Offering Direct-to-Consumer Genetic Testing

A total of 16 articles were identified regarding issues and marketing strategies related to the type of information provided by the DTC-GT websites (see Table 4 in Multimedia Appendix 2). The number of websites analyzed ranged from three [113] to 38 [114].

Goddard et al [115] found 27 health-related DTC-GT distributor websites and evaluated those that sold tests for thrombosis risk. Liu et al [116] analyzed 46 websites, but only 20 of them allowed consumers to order directly from the company. Sterling et al [117] identified 64 organizations hosting websites promoting nutrigenomic services, but only 29 offered or promoted at-home testing.

Borry et al [118] and Howard et al [119] investigated online companies focusing mainly on their policies in regard to GT for minors. The former analyzed 29 companies obtained from a list published by the Genetics and Policy Center, and the latter sent a questionnaire to 37 DTC-GT companies. Both studies emphasized a lack of exhaustive information on the privacy policy regarding minors, which is a deviation from the professional guidelines on this issue. Also, in a recent systematic Internet search for DTC genomic services, limited information on privacy policies was found [120]. This evidence contrasts with other studies that found the majority of sites selected provided this information [115-117,121,122].

Most studies assessed the quality of information provided by online GT companies through a content analysis of websites, with a focus on the provision of genetic counseling, suggestion for a physician's consultation, and the description of risks, benefits, and limitations of GTs.

In relation to online genetic counseling, Geransar et al [123] showed that of the 24 online companies studied, 75% recommended and arranged for counseling services. However, only one-third of the companies directly provided counseling services and just one of them provided a face-to-face format. Half of the websites analyzed by Covolo et al [124] provided this service pre- and/or post-test, with 20% offering this service for an extra fee. In other cases, fewer than 39% of online companies provided genetic counseling [114,119,121,125,126]. Pre-test counseling was rarely offered in studies conducted by Hennen et al [114], Lachance et al [121], and Liu et al [116]. None of the 29 companies offering nutrigenomic services examined by Sterling et al [117] provided genetic counseling.

Additionally, except for the websites analyzed by some studies [115,123,124,127], very few companies suggested a physician's consultation [114,117,119,121,125]. Sometimes the GT sale was accompanied by recommendations associated with disease prevention or health improvement (eg, nutritional supplements). This trend was found in the majority of websites (from 60-74%) investigated by Lewis et al [122] and Singleton et al [127]. Of 64 websites promoting nutrigenomic services identified by Sterling et al [117], 53% provided recommendations for dietary intake or supplementation.

Genetic discrimination, emotional consequences, risk of behavior changes, and confidentiality of test results are possible risks associated with GT. In general, all studies that searched for this information found that the risks were poorly cited, ranging from about 20% [115,116,122,128] to about 30-36% of the websites [117,124,125,127]. Of the company websites analyzed by Hennen et al [114], 47% provided information on consequences and actions to be taken in the case of a positive test result, and 37% in the case of a negative test result.

Clearly, the benefits of testing are described more than risks [115-117,124,127,128]. In particular, empowerment over one's health was highlighted by several authors [113,116,124,128]. Almost all of the sites identified by Lachance et al [121] and Singleton et al [127] listed at least one benefit to consumers by undergoing testing. Three-quarters (76%) of websites analyzed by Lachance et al [121] highlighted the fact that test results can help inform consumers in making a health decision. In the second study, prevention of the onset of a disease was the most common benefit presented (96%). Interestingly, 52% of websites stressed the consumer's ability to use the results to make informed decisions. The concept of patient empowerment also appeared in the Sterling et al [117] study. In fact, 73% (47/64) of organizations analyzed mentioned that consumers could use test results in their own diet and lifestyle decision making.

Over three-quarters (78%) of websites analyzed by Singleton et al [127] and about half of the websites analyzed by Lachance et al [121] and Lewis et al [122] mentioned limitations of test. None of the websites selling DTC-GT for thrombosis reported limitations [115].

Very little information or scientific evidence was provided on the clinical validity of tests [114-117,121-123,128]. Some websites referred to a laboratory certification, such as Clinical Laboratory Improvement Amendments (CLIA) standards, to indicate legitimacy [114,115,117,120-122,124,128].

Scientific Evidence and Clinical Utility of Direct-to-Consumer Genetic Testing

A total of 14 papers, including two reports from the US Government Accountability Office (GAO) [11,129], question the scientific quality, clinical validity, and utility of DTC-GT (see Table 1). This issue was addressed in different ways.

Table 1. List of articles on scientific evidence and clinical utility of direct-to-consumer genetic tests.

Author	Aim of the study	Main findings
Adams, 2013 [130]	To investigate the reliability and reproducibility of DTC-GT by sending DNA samples to 2 popular companies	DNA samples from 2 individuals were sent to both companies. For 5 of 14 health conditions for which both companies reported relative risk information, the results were conflicting. The significance of relative risk changes was overemphasized, given that they were associated with very small changes in absolute risk.
Bloss, 2012 [131]	To evaluate the relationship between DTC genomic risk estimates and self-reported disease of individuals who went on to purchase a DTC-GT	For 5 out of 15 total conditions studied, the risk estimates from the test were significantly associated with self-reported family and/or personal health history.
Buitendijk, 2014 [132]	To explore the practicability and predictive value of DTC tests from four companies for age-related macular degeneration in 3 individuals	Predicted risks varied widely within each individual, and differences between highest and lowest estimates for lifetime risk were up to 12-fold. Within the same person, overall relative risks could be increased as well as decreased, depending on which test was used. None may represent the true disease risk.
Imai, 2011 [133]	To evaluate 3 DTC services and genomics service and compare the test results obtained for the same individual	The concordance rates between the services for single nucleotide polymorphism (SNP) data were >99.6%. There were some marked differences in the relative disease risks assigned by the DTC services due to different SNPs used to calculate risk for the same disease.
Janssens, 2008 [7]	To assess the scientific evidence supporting the purported gene-associations for genes included in genomic profiles offered online	The seven companies investigated tested at least 69 polymorphisms in 56 genes. Of the 56 genes tested, 24 were not reviewed in meta-analyses. For the remaining 32 genes, they found 260 meta-analyses that examined 160 unique polymorphism-disease associations, of which only 60 were found to be statistically significant. However the associations were modest.
Johnson, 2010 [134]	To survey potential notifiable variants on arrays used in genome-wide association studies and DTC genetic services	They identified 298 specific targeted mutations, encompassing 56 disorders. Only 88 out of 298 mutations could be identified as known SNPs in genomic databases. Eighteen out of 88 SNPs were found in commercially available arrays.
Kalf, 2013 [135]	To examine and compare the methods of 3 companies offering DTC-GT	Predicted risks differed substantially among the companies as a result of differences in the sets of SNPs selected and the average population risks selected by the companies, and in the formulas used for the calculation of risks.
Kido, 2013 [136]	To evaluate the distributions of disease risk prediction from three DTC companies using three Japanese samples	The overall prediction results were correlated with each other, but not perfectly matched; less than one third mismatching of the opposite direction occurred in 8 diseases of 22.
Mihaescu, 2009 [137]	To investigate the extent to which updating of risk predictions from commercial genome-wide scans leads to reclassification of individuals from below to above average disease risk or vice versa taking type 2 diabetes as an example	At individual level, 34% of 5297 participants switched between risk categories when risks were updated from 1-18 polymorphisms and 29% switched when age, sex, and body mass index were considered. In total, 39% of participants switched risk categories once and 11% switched twice.
Ng, 2009 [138]	To compare results of tests purchased from two DTC companies on 13 diseases for 5 individuals	For seven diseases, 50% or less of the predictions of the two companies agreed across 5 individuals.
Palomaki, 2013 [139]	To review the evidence about the clinical and analytic validity of type 2 diabetes genomic risk profiles promulgated by DTC-GT companies	The quality of evidence for analytic validity was inadequate. Clinical validity ranged from inadequate to convincing for 30 variants identified on five T2D genomic panels. Clinical utility evidence was inadequate.
Swan, 2010 [140]	To understand the variance in risk interpretation for multigenic conditions among 5 genome-wide DTC genomic companies	Multigenic condition risk interpretation may vary between DTC genomic services due to differences in the average lifetime risk assigned to similar underlying populations, the loci and SNPs selected for analysis, and the quantitative risk assignment methodologies used by DTC genomic companies.
Kutz, 2006 [129]	To evaluate the results of nutrigenetic tests purchased from four DTC companies for 14 fictitious consumers coming from two DNA samples	All 14 results predicted risk of developing different medical conditions. These predictions were similar for all the fictitious consumers, no matter which DNA or lifestyle description they used. One of the four companies gave contradictory results.
Kutz, 2010 [11]	To compare results from 10 tests each purchased from four DTC companies on 15 diseases for 5 individuals. To assess whether the tests provided any medically useful information	Each donor received risk predictions for the 15 diseases that varied from company to company. Four of the five donors received test results that conflicted with their factual medical conditions and family histories.

Seven studies [11,129,130,132,133,136,138] focused on the comparison of GT results from DTC companies for one or more individuals. The first study, executed by the GAO in 2006 [129], evaluated the results of nutrigenetic tests purchased from four DTC companies for 14 fictitious consumers with different characteristics obtained from two DNA samples. Interestingly, all 14 results predicted the risk of developing different medical conditions. These predictions were similar for all of the fictitious consumers, no matter which DNA or lifestyle description was used. Only one of the four companies gave contradictory results.

In a more recent report by GAO [11], 5 individuals purchased 10 tests manufactured by four different DTC companies. The tests were specific to 15 diseases. The analysis found a large variation in prediction risk from company to company. In agreement with the GAO report, Ng et al [138] found a modest concordance among the results (50% or less) from two DTC companies on 13 diseases for 5 individuals.

Similarly, Imai et al [133] compared the relative common disease risks obtained from three DTC-GT companies for the same individual and found comparable results from the single nucleotide polymorphism (SNP) analyses from different companies. However, in a similar recent study [132], they also pointed out a large variation in relative risks for some of the diseases investigated, possibly due to different SNPs used to calculate the same disease, the choice of the reference population, and the risk calculation methodology.

Bloss et al [131] compared the DTC genomic risk estimates with self-reported disease from individuals who purchased a GT. The risk estimates were significantly associated with self-reported family or personal health history in only five out of 15 conditions studied. Two studies [135,140] examined the risk assessment of common diseases in DTC-genomic services and found that the predicted risks differed among the companies due to different methodologies used, different loci, and SNPs selected for analysis.

In an evaluation of type 2 diabetes risk prediction from commercial companies offering genome-wide scanning [137], it was shown that the individual risk prediction changed depending on the number of polymorphisms used to calculate the risk and characteristics of people (eg, age and gender). In particular, 39% of 5297 individuals switched between risk categories once and 11% switched twice. A study by Palomaki et al [139] of type 2 diabetes, genomic risk profiles advertised by DTC-GT companies highlighted a lack of analytical validity and clinical utility in the tests through the Evaluation of Genomic Applications in Practice and Prevention Working Group. This approach was established to support the development of a systematic process for assessing the available evidence for GT in clinical practice.

Other studies focused on the scientific evidence of genetic polymorphisms used to estimate the disease risk by DTC companies. In particular, Janssens et al [7] looked for meta-analyses supporting 69 polymorphisms tested by seven companies and found inconsequential scientific evidence. Similarly, it was found that only 18 out of 88 SNPs identified as known SNPs in genomic databases associated with a disease were present in a commercially available test [139].

Risk of Bias

The quality scores of the evidence reviewed ranged from 0.55-0.95. The majority of the studies that could be evaluated (69/80, 86%) had a score >0.7 (data not shown). Overall, all the studies are adequate in terms of methodological quality. The bias that was mostly present was a selection bias due to the recruitment of convenience population or a small sample size that did not allow a generalization of the results.

Discussion

Principal Findings

This review summarized the scientific literature on DTC-GT with a comprehensive view meant to unravel the complexity of the DTC-GT market. Previous systematic reviews dealt with this topic by focusing on certain aspects, particularly position statements, policies and recommendations [15], user perspectives [18], and health professionals’ perspectives [19]. This systematic review aimed to give an overall view of the DTC-GT market to include studies that analyzed the content of the websites offering these products, as well as studies focused on the scientific evidence and clinical utility of such tests. The large number of reports retrieved on this issue indicates a strong interest in the topic.

Thanks to the prevalence of the Internet over the past decade, the availability of health-related products on a DTC basis has become increasingly common. However, the fact that the promotion of these products such as drugs [141] or nutritional supplements [142] is comparable to the sale of any commodity is a cause for concern.

In terms of marketing, we must discuss the results considering the product, the offer, and the potential customers, in addition to the opinion of health professionals as product experts.

The Product

Advances in genomic technology made GT available for both monogenic disorders and common complex diseases, in addition to nutrigenetic and pharmacogenetic tests. To date, the majority of these tests have provided a poor predictive value, and the assessment of the clinical validity and utility is still a work in progress. However, many commercial companies have begun to bring these tests to market. Their lack of scientific evidence was confirmed by some studies focused particularly on commercially available GTs [7,134]. Overall, all studies comparing the results of GT of the same people from different companies showed a modest concordance in risk predictions and sometimes reported contrasting results [11,129,130,132,133,136,138].

The Offer

In general, the studies focused on the content of websites selling DTC-GT agree that these companies do not provide complete information to the consumer. This emphasizes the poor quality of information on the scientific evidence and clinical validity to support the tests. It also highlights the lack of attention to the risks related to the performance of the tests, genetic discrimination, emotional consequences, behavior changes, and confidentiality of test results. Furthermore, genetic counseling requirements were often missing. As with other sales-oriented companies, these websites contain marketing strategies that accentuate the benefits of the product. The benefits of testing were described more than the risks, and the theme of patient empowerment is highly emphasized as a good reason for testing [113,116,124,128]. The main emphasis on genes, without consideration for environment, might lead consumers to misinterpret test results, as was found by the majority of studies that addressed this issue.

The Customer

Thirty-seven papers examined consumer knowledge of and attitude toward DTC-GT; likewise 37 papers discussed the impact of these tests on users. In general, it was confirmed that consumers have an interest in DTC-GT and that their main motivation is curiosity, as well as some interest in monitoring and improving health.

It should be noted that interest is highest among employees of biotechnology companies [30] or people at risk for cancer [35] and other diseases [25]—not the general public. Additionally, study participants were highly educated.

As for the impact of DTC-GT on users, nine papers and eight case studies were reviewed, in addition to those researched by Goldsmith et al [18]. The research remains limited because it evaluates the actual consequences of having a DTC test. Other studies reported situations where participants were exposed to mock reports about their genetic susceptibility or were asked to voluntarily submit to testing for research purposes. It seems that the feared negative consequences, such as psychological impact or increased anxiety for consumers, were not confirmed. Similarly, positive consequences, such as adoption of healthier lifestyle behaviors, were not observed, although a large proportion of people expressed the intention to change lifestyles [43,55,63,80].

Negative consequences may arise from misinterpretation of test results, which is another aspect addressed by several studies. The majority of studies showed that participants did not have particular difficulty understanding the test results. In fact, only a small number of people shared their results with a physician and reported worry after receiving them. Yet, it was also determined that the presence of a professional provided better interpretations of results compared to participants who received results online [87]. Interestingly, incorrect interpretation of results was confirmed in actual DTC-GT customers [43,91].

The Expert

We considered health professionals' perspectives as expert opinions, considering the strong recommendation for involvement of a health professional in the order process and interpretation of test results [143]. Compared to previous systematic reviews [19], 12 additional articles were retrieved that focused on health professionals' perspectives and the overall scenario described by these authors were confirmed by our findings. As stated by Goldsmith et al [19], the level of awareness of DTC-GT remains inconsistent, even with three studies [41,100,110] in which the majority of physicians are aware of DTC-GT (but the sample size was 50 participants). In addition, few respondents have had direct experience with DTC-GT. The overall opinions regarding the utility of the tests are contrasting. Some professionals are in favor of some GT [99,101,102], while others considered it unacceptable to provide a predictive genetic test without genetic counseling [110] and were concerned about possible psychological harm [107], misunderstanding of results, and insurance discrimination [98]. Understandably, clinical geneticists expressed more concerns than general practitioners [107,110,111]. It should be noted that few physicians considered themselves confident in interpreting GT results and reported the need for more training [97,106,107,109]. In fact, an increase in the incorporation of instruction about application and technique in predictive genomic testing was presented in a recent study [103].

As discussed by some authors [97,144], these concerns raise the question of whether a non-geneticist physician involved in the commercial distribution of GT is properly equipped to offer test information to patients. So the presence of a physician does not guarantee the provision of adequate information. This is a cause for concern considering the recent shift from selling tests directly to the consumer to a direct-to-provider marketing model [97,118,144].

Implications for Policy Making

We believe this review highlights the important aspects in considering the regulation of DTC-GT from a policy perspective. More specifically, there are at least three main issues to address to improve DTC-GT for a better service for the public:

1. DTC-GT is currently advertised despite the minimal and controversial nature of the supporting evidence. Here, more research is needed to evaluate these products and to eventually decide whether or not it is appropriate to market them at all.
2. As for other DTC products, GT is advertised by means of traditional strategies of persuasion generally used for commercial products (eg, more emphasis on benefits than on side effects). The rhetorical selling of DTC products calls for an enrichment of the guidelines for advertising of health-related products. In particular, these guidelines need to take into consideration the important literature from the fields of rhetoric and persuasion that explain how communication can be used to manipulate the beliefs and attitudes of consumers. The marketing of DTC-GT cannot be biased as it currently appears from the content analysis of websites.
3. From an ethical point of view, the first question to answer is whether, in light of the limitations in evidence and communication, DTC-GT empowers consumers. If empowerment is valuable because it is linked to autonomy, does current DTC-GT contribute beneficially to the development and application of autonomy? A second question concerns the fact that DTC-GT promotes products whose social implications have not been properly addressed. How does knowledge of self-assessed genetic risks influence the life of consumers? Current marketing of GT seems to be mainly interested in the advantages, as advantages convince consumers to purchase. But empowerment cannot be promoted separately from a full

> appraisal of the ethical aspects surrounding the delivery of a specific type of information.

Overall, research and practice must collaborate toward policy making in a field that is already open to the public despite its serious pitfalls.

Limitations

Through a systematic approach we aimed to provide a comprehensive look at the DTC-GT market in order to better understand its actual impact on population. Although the number of articles retrieved is relatively large, some limitations related to studies design should be underlined. The majority of the studies used a cross-sectional design. It is known that response rate as well as sample size and sample selection are critical points in this kind of design [145]. Considering the total of the surveys investigating the awareness, use, and perceptions of health professionals and consumers, only 9 studies [30-33,40,52,56,59,81,92] out of 56 surveys (16%) (see Multimedia Appendix 2) have a response rate of at least 50% and a large sample size (more than 1000 subjects).

Most of the subjects recruited were highly educated and sometimes selection bias was present (eg, employees of health and technology companies [81]). All these aspects mean a poor representativeness of population [145]. Furthermore, few health professionals and consumers had direct experience with DTC-GT, so as previously argued [18,19], the responses of participants based on hypothetical scenarios make it difficult to draw conclusions about the actual impact of DTC-GT market.

Conclusions

Based on the evidence collected, it seems that DTC-GT is neither beneficial nor detrimental to potential users. It should also be noted that the development of online companies is rapidly changing, most likely due to pressure from government agencies such as the FDA. Some companies have also changed their delivery model to include the health profession in the order process [143].

However, regardless of the large amount of data available on this issue, the actual experiences of DTC-GT users are still limited and this market is still in the early stages of distribution to the general public. Furthermore, some limitations on previous studies must be addressed. For instance, the majority of studies are characterized by people who do not represent the general public (participants were often convenience samples), featuring low sample size or limited response rate. Additionally, the prospective studies typically employ relatively short-term follow-up in the majority of the cases, not sufficient to evaluate the impact of DTC-GT on behavioral changes.

On the other hand, it is unacceptable that online companies offer GT lacking scientific evidence, no proven clinical utility, and misleading marketing claims. As underscored by Janssens and van Duijn [146], the expected benefits of whole genome scanning may be larger when tests are targeted only to specific at-risk populations, and not to populations-at-large, because of the moderate predictive ability of these current tests.

According to global industry analysts, the global genetic testing market is expected to reach more than US $230 million by 2018 [147]. Combined with the rapid decrease in biotechnology costs, this revenue stream will eventually allocate testing accessibility to all socioeconomic classes. It is important, therefore, to remain cautious and vigilant about this growing, influential health care market.

Conflicts of Interest

None declared.

Multimedia Appendix 1

PRISMA checklist.

[PDF File (Adobe PDF File), 150KB - jmir_v17i12e279_app1.pdf]

Multimedia Appendix 2

Supplementary tables.

[PDF File (Adobe PDF File), 425KB - jmir_v17i12e279_app2.pdf]

References

1. Niccol A. Gattaca. Los Angeles, CA: Columbia pictures; 1997. URL: http://www.imdb.com/title/tt0119177/quotes [WebCite Cache ID 6d6WcaXRO]
2. easy-DNA. URL: https://www.easy-dna.com/genetic-predisposition-dna-testing/ [accessed 2015-11-16] [WebCite Cache ID 6d51iF0xK]
3. Test Country. Test Country. URL: http://www.testcountry.com/products.html?product=1966 [accessed 2015-02-12] [WebCite Cache ID 6WI0DmVaE]
4. Inherent Health. URL: http://www.inherenthealth.com/ [accessed 2015-02-12] [WebCite Cache ID 6WI0MYcrT]
5. Gene Partner. URL: http://www.genepartner.com/ [accessed 2015-02-12] [WebCite Cache ID 6WI01otjv]

6. Cornel MC, van El CG, Borry P. The challenge of implementing genetic tests with clinical utility while avoiding unsound applications. J Community Genet 2014 Jan;5(1):7-12 [FREE Full text] [doi: 10.1007/s12687-012-0121-1] [Medline: 23055102]
7. Janssens CW, Gwinn M, Bradley LA, Oostra BA, van Duijn CM, Khoury MJ. A critical appraisal of the scientific basis of commercial genomic profiles used to assess health risks and personalize health interventions. Am J Hum Genet 2008 Mar;82(3):593-599 [FREE Full text] [doi: 10.1016/j.ajhg.2007.12.020] [Medline: 18319070]
8. Berg C, Fryer-Edwards K. The Ethical Challenges of Direct-to-Consumer Genetic Testing. J Bus Ethics 2008;77:17-31.
9. Borry P, Cornel MC, Howard HC. Where are you going, where have you been: a recent history of the direct-to-consumer genetic testing market. J Community Genet 2010 Sep;1(3):101-106 [FREE Full text] [doi: 10.1007/s12687-010-0023-z] [Medline: 21475669]
10. Hogarth S. Myths, misconceptions and myopia: searching for clarity in the debate about the regulation of consumer genetics. Public Health Genomics 2010;13(5):322-326 [FREE Full text] [doi: 10.1159/000313330] [Medline: 20588076]
11. Kutz G. Direct-to-consumer genetic tests: Misleading test results are further complicated by deceptive marketing and other questionable practices.: United States Government Accountability Office URL: http://www.gao.gov/assets/130/125079.pdf [accessed 2015-02-12] [WebCite Cache ID 6WI0aktrv]
12. Edwards KT, Huang CJ. Bridging the consumer-medical divide: how to regulate direct-to-consumer genetic testing. Hastings Cent Rep 2014;44(3):17-19. [doi: 10.1002/hast.310] [Medline: 24821248]
13. Hudson K, Javitt G, Burke W, Byers P, American Society of Human Genetics Social Issues Committee. ASHG Statement* on direct-to-consumer genetic testing in the United States. Obstet Gynecol 2007 Dec;110(6):1392-1395. [doi: 10.1097/01.AOG.0000292086.98514.8b] [Medline: 18055737]
14. European Academies Science Advisory Council (EASAC)/Federation of European Academies of Medicine. EASAC policy report 18. 2012. Direct-to-consumer genetic testing for health-related purposes in the European Union URL: http://www.easac.eu/fileadmin/Reports/EASAC_Genetic_Testing_Web_complete.pdf [accessed 2015-02-12] [WebCite Cache ID 6WI0i2URW]
15. Skirton H, Goldsmith L, Jackson L, O'Connor A. Direct to consumer genetic testing: a systematic review of position statements, policies and recommendations. Clin Genet 2012 Sep;82(3):210-218. [doi: 10.1111/j.1399-0004.2012.01863.x] [Medline: 22372614]
16. Public Health Service Food and Drug Administration. Warning Letter. 2013 Nov 22. URL: http://www.fda.gov/iceci/enforcementactions/warningletters/2013/ucm376296.htm [WebCite Cache ID 6WI0tSiDC]
17. GenePlanet. URL: http://www.geneplanet.com/ [accessed 2015-02-12] [WebCite Cache ID 6WI0yfBfI]
18. Goldsmith L, Jackson L, O'Connor A, Skirton H. Direct-to-consumer genomic testing: systematic review of the literature on user perspectives. Eur J Hum Genet 2012 Aug;20(8):811-816 [FREE Full text] [doi: 10.1038/ejhg.2012.18] [Medline: 22333900]
19. Goldsmith L, Jackson L, O'Connor A, Skirton H. Direct-to-consumer genomic testing from the perspective of the health professional: a systematic review of the literature. J Community Genet 2013 Apr;4(2):169-180 [FREE Full text] [doi: 10.1007/s12687-012-0135-8] [Medline: 23322235]
20. Roberts JS, Ostergren J. Direct-to-Consumer Genetic Testing and Personal Genomics Services: A Review of Recent Empirical Studies. Curr Genet Med Rep 2013 Sep;1(3):182-200 [FREE Full text] [Medline: 24058877]
21. Saukko P. State of play in direct-to-consumer genetic testing for lifestyle-related diseases: market, marketing content, user experiences and regulation. Proc Nutr Soc 2013 Feb;72(1):53-60 [FREE Full text] [doi: 10.1017/S0029665112002960] [Medline: 23336560]
22. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
23. Wikipedia. PageRank 2014. URL: http://en.wikipedia.org/wiki/PageRank [WebCite Cache ID 6WI18UKJM]
24. Kmet L, Lee R, Cook L. Alberta Heritage Foundation for Medical Research. Edmonton, Canada; 2004 Feb 01. Standard Quality Assessment criteria for Evaluating Primary Research Papers from a Variety of Fields URL: http://www.ihe.ca/advanced-search?date=2004&type=1021Archived [WebCite Cache ID 6bxHoayCU]
25. Wilde A, Meiser B, Mitchell PB, Schofield PR. Public interest in predictive genetic testing, including direct-to-consumer testing, for susceptibility to major depression: preliminary findings. Eur J Hum Genet 2010 Jan;18(1):47-51 [FREE Full text] [doi: 10.1038/ejhg.2009.138] [Medline: 19690586]
26. Lynch J, Parrott A, Hopkin RJ, Myers M. Media coverage of direct-to-consumer genetic testing. J Genet Couns 2011 Oct;20(5):486-494. [doi: 10.1007/s10897-011-9374-9] [Medline: 21638197]
27. Rahm AK, Feigelson HS, Wagner N, Le AQ, Halterman E, Cornish N, et al. Perception of direct-to-consumer genetic testing and direct-to-consumer advertising of genetic tests among members of a large managed care organization. J Genet Couns 2012 Jun;21(3):448-461. [doi: 10.1007/s10897-011-9477-3] [Medline: 22278218]
28. Wasson K, Hogan NS, Sanders TN, Helzlsouer KJ. Primary Care Patients’ Views, Attitudes, and Decision-Making Factors Regarding Direct-to-Consumer Personal Genome Testing: Results From a Qualitative Study. AJOB Primary Research 2012 Apr;3(2):24-35. [doi: 10.1080/21507716.2011.650344]

29. Adams HR, Rose K, Augustine EF, Kwon JM, deBlieck EA, Marshall FJ, et al. Experience, knowledge, and opinions about childhood genetic testing in Batten disease. Mol Genet Metab 2014 Feb;111(2):197-202 [FREE Full text] [doi: 10.1016/j.ymgme.2013.10.017] [Medline: 24246680]
30. Bloss CS, Ornowski L, Silver E, Cargill M, Vanier V, Schork NJ, et al. Consumer perceptions of direct-to-consumer personalized genomic risk assessments. Genet Med 2010 Sep;12(9):556-566. [doi: 10.1097/GIM.0b013e3181eb51c6] [Medline: 20717041]
31. Finney RL, Gollust S, Naveed S, Moser R. Increasing Public Awareness of Direct-to-Consumer Genetic Tests: Health Care Access, Internet Use, and Population Density Correlates. J Cancer Epidemiol 2012 Jul 30:1-7. [doi: 10.1155/2012/309109]
32. Goddard K, Moore C, Ottman D, Szegda KL, Bradley L, Khoury MJ. Awareness and use of direct-to-consumer nutrigenomic tests, United States, 2006. Genet Med 2007 Aug;9(8):510-517. [Medline: 17700389]
33. Goddard K, Duquette D, Zlot A, Johnson J, Annis-Emeott A, Lee PW, et al. Public awareness and use of direct-to-consumer genetic tests: results from 3 state population-based surveys, 2006. Am J Public Health 2009 Mar;99(3):442-445. [doi: 10.2105/AJPH.2007.131631] [Medline: 19106425]
34. Gollust SE, Gordon ES, Zayac C, Griffin G, Christman MF, Pyeritz RE, et al. Motivations and perceptions of early adopters of personalized genomics: perspectives from research participants. Public Health Genomics 2012;15(1):22-30 [FREE Full text] [doi: 10.1159/000327296] [Medline: 21654153]
35. Graves KD, Peshkin BN, Luta G, Tuong W, Schwartz MD. Interest in genetic testing for modest changes in breast cancer risk: implications for SNP testing. Public Health Genomics 2011;14(3):178-189 [FREE Full text] [doi: 10.1159/000324703] [Medline: 21464556]
36. Gray SW, Hornik RC, Schwartz JS, Armstrong K. The impact of risk information exposure on women's beliefs about direct-to-consumer genetic testing for BRCA mutations. Clin Genet 2012 Jan;81(1):29-37 [FREE Full text] [doi: 10.1111/j.1399-0004.2011.01797.x] [Medline: 21992449]
37. Hall TO, Renz AD, Snapinn KW, Bowen DJ, Edwards KL. Awareness and uptake of direct-to-consumer genetic testing among cancer cases, their relatives, and controls: the Northwest Cancer Genetics Network. Genet Test Mol Biomarkers 2012 Jul;16(7):744-748 [FREE Full text] [doi: 10.1089/gtmb.2011.0235] [Medline: 22731649]
38. Kaphingst KA, McBride CM, Wade C, Alford SH, Brody LC, Baxevanis AD. Consumers' use of web-based information and their decisions about multiplex genetic susceptibility testing. J Med Internet Res 2010;12(3):e41 [FREE Full text] [doi: 10.2196/jmir.1587] [Medline: 20884465]
39. Kolor K, Duquette D, Zlot A, Foland J, Anderson B, Giles R, et al. Public awareness and use of direct-to-consumer personal genomic tests from four state population-based surveys, and implications for clinical and public health practice. Genet Med 2012 Oct;14(10):860-867. [doi: 10.1038/gim.2012.67] [Medline: 22814860]
40. Langford AT, Resnicow K, Roberts JS, Zikmund-Fisher BJ. Racial and ethnic differences in direct-to-consumer genetic tests awareness in HINTS 2007: sociodemographic and numeracy correlates. J Genet Couns 2012 Jun;21(3):440-447. [doi: 10.1007/s10897-011-9478-2] [Medline: 22271378]
41. Leighton JW, Valverde K, Bernhardt BA. The general public's understanding and perception of direct-to-consumer genetic test results. Public Health Genomics 2012;15(1):11-21 [FREE Full text] [doi: 10.1159/000327159] [Medline: 21720148]
42. McBride CM, Alford SH, Reid RJ, Larson EB, Baxevanis AD, Brody LC. Characteristics of users of online personalized genomic risk assessments: implications for physician-patient interactions. Genet Med 2009 Aug;11(8):582-587 [FREE Full text] [doi: 10.1097/GIM.0b013e3181b22c3a] [Medline: 19606049]
43. McGuire AL, Diaz CM, Wang T, Hilsenbeck SG. Social networkers' attitudes toward direct-to-consumer personal genome testing. Am J Bioeth 2009;9(6-7):3-10 [FREE Full text] [doi: 10.1080/15265160902928209] [Medline: 19998099]
44. O'Neill S, White D, Sanderson S, Lipkus I, Bepler G, Bastian L, et al. The feasibility of online genetic testing for lung cancer susceptibility: uptake of a web-based protocol and decision outcomes. Genet Med 2008 Feb;10(2):121-130.
45. Ortiz AP, López M, Flores LT, Soto-Salgado M, Finney Rutten LJ, Serrano-Rodriguez RA, et al. Awareness of direct-to-consumer genetic tests and use of genetic tests among Puerto Rican adults, 2009. Prev Chronic Dis 2011 Sep;8(5):A110 [FREE Full text] [Medline: 21843413]
46. Paquin RS, Richards AS, Koehly LM, McBride CM. Exploring dispositional tendencies to seek online information about direct-to-consumer genetic testing. Transl Behav Med 2012 Dec;2(4):392-400 [FREE Full text] [doi: 10.1007/s13142-012-0159-y] [Medline: 24073142]
47. Perez GK, Cruess DG, Cruess S, Brewer M, Stroop J, Schwartz R, et al. Attitudes toward direct-to-consumer advertisements and online genetic testing among high-risk women participating in a hereditary cancer clinic. J Health Commun 2011 Jul;16(6):607-628. [doi: 10.1080/10810730.2011.551993] [Medline: 21432710]
48. Ramakrishnan S, Rollins B, Perri MI. Direct-to-consumer advertising of predictive genetic tests: an examination of consumer attitudes, behavioral intentions and information seeking behavior. J Health & Mass Comm 2011;3:1940-1943.
49. Sweeny K, Legg AM. Predictors of interest in direct-to-consumer genetic testing. Psychol Health 2011 Oct;26(10):1259-1272. [doi: 10.1080/08870446.2010.514607] [Medline: 21500101]
50. Tercyak KP, Hensley AS, Emmons KM, Lipkus IM, Wilfond BS, McBride CM. Parents' attitudes toward pediatric genetic testing for common disease risk. Pediatrics 2011 May;127(5):e1288-e1295 [FREE Full text] [doi: 10.1542/peds.2010-0938] [Medline: 21502235]

51. Critchley C, Nicol D, Otlowski M, Chalmers D. Public reaction to direct-to-consumer online genetic tests: Comparing attitudes, trust and intentions across commercial and conventional providers. Public Underst Sci 2015 Aug;24(6):731-750.
52. Mai Y, Koromila T, Sagia A, Cooper DN, Vlachopoulos G, Lagoumintzis G, et al. A critical view of the general public’s awareness and physicians’ opinion of the trends and potential pitfalls of genetic testing in Greece. Personalized Medicine 2011 Sep;8(5):551-561. [doi: 10.2217/pme.11.48]
53. Ries NM, Hyde-Lay R, Caulfield T. Willingness to pay for genetic testing: a study of attitudes in a Canadian population. Public Health Genomics 2010;13(5):292-300. [doi: 10.1159/000253120] [Medline: 19864872]
54. Sanderson SC, Wardle J. Associations between anticipated reactions to genetic test results and interest in genetic testing: will self-selection reduce the potential for harm? Genet Test 2008 Mar;12(1):59-66. [doi: 10.1089/gte.2007.0047] [Medline: 18318647]
55. Vayena E, Ineichen C, Stoupka E, Hafen E. Playing a part in research? University students' attitudes to direct-to-consumer genomics. Public Health Genomics 2014;17(3):158-168 [FREE Full text] [doi: 10.1159/000360257] [Medline: 24777115]
56. Wilde A, Meiser B, Mitchell PB, Hadzi-Pavlovic D, Schofield PR. Community interest in predictive genetic testing for susceptibility to major depressive disorder in a large national sample. Psychol Med 2011 Aug;41(8):1605-1613. [doi: 10.1017/S0033291710002394] [Medline: 21159212]
57. Nielsen DE, El-Sohemy A. A randomized trial of genetic information for personalized nutrition. Genes Nutr 2012 Oct;7(4):559-566 [FREE Full text] [doi: 10.1007/s12263-012-0290-x] [Medline: 22407352]
58. Savard J, Mooney-Somers J, Newson AJ, Kerridge I. Australians' knowledge and perceptions of direct-to-consumer personal genome testing. Intern Med J 2014 Jan;44(1):27-31. [doi: 10.1111/imj.12289] [Medline: 24450520]
59. Cherkas LF, Harris JM, Levinson E, Spector TD, Prainsack B. A survey of UK public interest in internet-based personal genome testing. PLoS One 2010;5(10):e13473 [FREE Full text] [doi: 10.1371/journal.pone.0013473] [Medline: 20976053]
60. Vermeulen E, Henneman L, van El CG, Cornel MC. Public attitudes towards preventive genomics and personal interest in genetic testing to prevent disease: a survey study. Eur J Public Health 2014 Oct;24(5):768-775 [FREE Full text] [doi: 10.1093/eurpub/ckt143] [Medline: 24068545]
61. Rollins BL, Ramakrishnan S, Perri M. Direct-to-consumer advertising of predictive genetic tests: a health belief model based examination of consumer response. Health Mark Q 2014;31(3):263-278. [doi: 10.1080/07359683.2014.936295] [Medline: 25120046]
62. Harris A, Kelly SE, Wyatt S. Autobiologies on YouTube: Narratives of Direct-to-Consumer Genetic Testing. New Genet Soc 2014 Mar;33(1):60-78. [doi: 10.1080/14636778.2014.884456] [Medline: 24772003]
63. Nusbaum R, Leventhal K, Hooker GW, Peshkin BN, Butrick M, Salehizadeh Y, et al. Translational Genomic Research: Protocol Development and Initial Outcomes following SNP Testing for Colon Cancer Risk. Transl Behav Med 2013 Mar 1;3(1):17-29 [FREE Full text] [doi: 10.1007/s13142-012-0149-0] [Medline: 23565131]
64. Vernez SL, Salari K, Ormond KE, Lee SS. Personal genome testing in medical education: student experiences with genotyping in the classroom. Genome Med 2013;5(3):24 [FREE Full text] [doi: 10.1186/gm428] [Medline: 23510111]
65. Wasson K, Sanders TN, Hogan NS, Cherny S, Helzlsouer KJ. Primary care patients' views and decisions about, experience of and reactions to direct-to-consumer genetic testing: a longitudinal study. J Community Genet 2013 Oct;4(4):495-505 [FREE Full text] [doi: 10.1007/s12687-013-0156-y] [Medline: 23832288]
66. Wasson K, Cherny S, Sanders TN, Hogan NS, Helzlsouer KJ. Who are you going to call? Primary care patients' disclosure decisions regarding direct-to-consumer genetic testing. Narrat Inq Bioeth 2014;4(1):53-68. [doi: 10.1353/nib.2014.0026] [Medline: 24748260]
67. McGowan ML, Fishman JR, Lambrix MA. Personal genomics and individual identities: motivations and moral imperatives of early users. New Genet Soc 2010 Sep 1;29(3):261-290 [FREE Full text] [doi: 10.1080/14636778.2010.507485] [Medline: 21076647]
68. Su Y, Howard HC, Borry P. Users' motivations to purchase direct-to-consumer genome-wide testing: an exploratory study of personal stories. J Community Genet 2011 Sep;2(3):135-146 [FREE Full text] [doi: 10.1007/s12687-011-0048-y] [Medline: 22109820]
69. Su Y, Borry P, Otte IC, Howard HC. “It’s our DNA, we deserve the right to test!” A content analysis of a petition for the right to access direct-to-consumer genetic testing. Personalized Medicine 2013 Sep;10(7):729-739. [doi: 10.2217/pme.13.69]
70. Gordon ES, Griffin G, Wawak L, Pang H, Gollust SE, Bernhardt BA. "It's not like judgment day": public understanding of and reactions to personalized genomic risk information. J Genet Couns 2012 Jun;21(3):423-432 [FREE Full text] [doi: 10.1007/s10897-011-9476-4] [Medline: 22180182]
71. Vayena E, Gourna E, Streuli J, Hafen E, Prainsack B. Experiences of early users of direct-to-consumer genomics in Switzerland: an exploratory study. Public Health Genomics 2012;15(6):352-362. [doi: 10.1159/000343792] [Medline: 23154382]
72. Austin SEA, Hegele RA. Clinical implications of direct-to-consumer genetic testing for cardiovascular disease risk. Can J Cardiol 2011;27(6):682-684. [doi: 10.1016/j.cjca.2011.02.006] [Medline: 21652171]
73. Brownstein CA, Margulies DM, Manzi SF. Misinterpretation of TPMT by a DTC genetic testing company. Clin Pharmacol Ther 2014 Jun;95(6):598-600. [doi: 10.1038/clpt.2014.60] [Medline: 24714787]

74. Corpas M. A family experience of personal genomics. J Genet Couns 2012 Jun;21(3):386-391. [doi: 10.1007/s10897-011-9473-7] [Medline: 22223063]
75. Dohany L, Gustafson S, Ducaine W, Zakalik D. Psychological distress with direct-to-consumer genetic testing: a case report of an unexpected BRCA positive test result. J Genet Couns 2012 Jun;21(3):399-401. [doi: 10.1007/s10897-011-9475-5] [Medline: 22271377]
76. Messner DA. Informed Choice in Direct-to-Consumer Genetic Testing for Alzheimer and Other Diseases: Lessons from Two Cases. New Genet Soc 2011;30(1):59-72 [FREE Full text] [doi: 10.1080/14636778.2011.552300] [Medline: 21603253]
77. Roberts ME, Riegert-Johnson DL, Thomas BC. Self diagnosis of Lynch syndrome using direct to consumer genetic testing: a case study. J Genet Couns 2011 Aug;20(4):327-329. [doi: 10.1007/s10897-011-9356-y] [Medline: 21448764]
78. Saunders KH, Nazareth S, Pressman PI. Case report: BRCA in the Ashkenazi population: are current testing guidelines too exclusive? Hered Cancer Clin Pract 2011;9(1):3 [FREE Full text] [doi: 10.1186/1897-4287-9-3] [Medline: 21711529]
79. Sturm AC, Manickam K. Direct-to-consumer personal genomic testing: a case study and practical recommendations for “genomic counseling”. J Genet Couns 2012 Jun;21(3):402-412. [doi: 10.1007/s10897-012-9489-7] [Medline: 22302621]
80. Bansback N, Sizto S, Guh D, Anis AH. The effect of direct-to-consumer genetic tests on anticipated affect and health-seeking behaviors: a pilot survey. Genet Test Mol Biomarkers 2012 Oct;16(10):1165-1171. [doi: 10.1089/gtmb.2012.0074] [Medline: 22931244]
81. Bloss CS, Schork NJ, Topol EJ. Effect of direct-to-consumer genomewide profiling to assess disease risk. N Engl J Med 2011 Feb 10;364(6):524-534 [FREE Full text] [doi: 10.1056/NEJMoa1011893] [Medline: 21226570]
82. Bloss CS, Wineinger NE, Darst BF, Schork NJ, Topol EJ. Impact of direct-to-consumer genomic testing at long term follow-up. J Med Genet 2013 Jun;50(6):393-400. [doi: 10.1136/jmedgenet-2012-101207] [Medline: 23559530]
83. Bloss CS, Schork NJ, Topol EJ. Direct-to-consumer pharmacogenomic testing is associated with increased physician utilisation. J Med Genet 2014 Feb;51(2):83-89. [doi: 10.1136/jmedgenet-2013-101909] [Medline: 24343916]
84. Dar-Nimrod I, Zuckerman M, Duberstein PR. The effects of learning about one's own genetic susceptibility to alcoholism: a randomized experiment. Genet Med 2013 Feb;15(2):132-138. [doi: 10.1038/gim.2012.111] [Medline: 22935722]
85. Darst BF, Madlensky L, Schork NJ, Topol EJ, Bloss CS. Perceptions of genetic counseling services in direct-to-consumer personal genomic testing. Clin Genet 2013 Oct;84(4):335-339. [doi: 10.1111/cge.12166] [Medline: 23590221]
86. Darst BF, Madlensky L, Schork NJ, Topol EJ, Bloss CS. Characteristics of genomic test consumers who spontaneously share results with their health care provider. Health Commun 2014;29(1):105-108 [FREE Full text] [doi: 10.1080/10410236.2012.717216] [Medline: 23384116]
87. Haga SB, Barry WT, Mills R, Svetkey L, Suchindran S, Willard HF, et al. Impact of delivery models on understanding genomic risk for type 2 diabetes. Public Health Genomics 2014;17(2):95-104 [FREE Full text] [doi: 10.1159/000358413] [Medline: 24577154]
88. James KM, Cowl CT, Tilburt JC, Sinicrope PS, Robinson ME, Frimannsdottir KR, et al. Impact of direct-to-consumer predictive genomic testing on risk perception and worry among patients receiving routine care in a preventive health clinic. Mayo Clin Proc 2011 Oct;86(10):933-940 [FREE Full text] [doi: 10.4065/mcp.2011.0190] [Medline: 21964170]
89. Kaphingst KA, McBride CM, Wade C, Alford SH, Reid R, Larson E, et al. Patients' understanding of and responses to multiplex genetic susceptibility test results. Genet Med 2012 Jul;14(7):681-687 [FREE Full text] [doi: 10.1038/gim.2012.22] [Medline: 22481132]
90. Francke U, Dijamco C, Kiefer A, Eriksson N, Moiseff B, Tung J, et al. Dealing with the unexpected: consumer responses to direct-access BRCA mutation testing. PeerJ 2013 Feb:1-24. [doi: 10.7717/peerj.8]
91. Kaufman DJ, Bollinger JM, Dvoskin RL, Scott JA. Risky business: risk perception and the use of medical services among customers of DTC personal genetic testing. J Genet Couns 2012 Jun;21(3):413-422. [doi: 10.1007/s10897-012-9483-0] [Medline: 22278220]
92. Reid RJ, McBride CM, Alford SH, Price C, Baxevanis AD, Brody LC, et al. Association between health-service use and multiplex genetic testing. Genet Med 2012 Oct;14(10):852-859 [FREE Full text] [doi: 10.1038/gim.2012.52] [Medline: 22595941]
93. Boeldt DL, Schork NJ, Topol EJ, Bloss CS. Influence of individual differences in disease perception on consumer response to direct-to-consumer genomic testing. Clin Genet 2015 Mar;87(3):225-232. [doi: 10.1111/cge.12419] [Medline: 24798746]
94. Egglestone C, Morris A, O'Brien A. Effect of direct-to-consumer genetic tests on health behaviour and anxiety: a survey of consumers and potential consumers. J Genet Couns 2013 Oct;22(5):565-575. [doi: 10.1007/s10897-013-9582-6] [Medline: 23547022]
95. Nielsen DE, Shih S, El-Sohemy A. Perceptions of genetic testing for personalized nutrition: a randomized trial of DNA-based dietary advice. J Nutrigenet Nutrigenomics 2014;7(2):94-104. [doi: 10.1159/000365508] [Medline: 25170902]
96. Weir M, Morin K, Ries N, Castle D. Canadian health care professionals' knowledge, attitudes and perceptions of nutritional genomics. Br J Nutr 2010 Oct;104(8):1112-1119. [doi: 10.1017/S0007114510002035] [Medline: 20550743]
97. McGowan ML, Fishman JR, Settersten RA, Lambrix MA, Juengst ET. Gatekeepers or intermediaries? The role of clinicians in commercial genomic testing. PLoS One 2014;9(9):e108484 [FREE Full text] [doi: 10.1371/journal.pone.0108484] [Medline: 25259512]

98. Bernhardt BA, Zayac C, Gordon ES, Wawak L, Pyeritz RE, Gollust SE. Incorporating direct-to-consumer genomic information into patient care: attitudes and experiences of primary care physicians. Per Med 2012 Sep 1;9(7):683-692 [FREE Full text] [doi: 10.2217/pme.12.80] [Medline: 23795206]
99. Giovanni MA, Fickie MR, Lehmann LS, Green RC, Meckley LM, Veenstra D, et al. Health-care referrals from direct-to-consumer genetic testing. Genet Test Mol Biomarkers 2010 Dec;14(6):817-819 [FREE Full text] [doi: 10.1089/gtmb.2010.0051] [Medline: 20979566]
100. Haga SB, Carrig MM, O'Daniel JM, Orlando LA, Killeya-Jones LA, Ginsburg GS, et al. Genomic risk profiling: attitudes and use in personal and clinical care of primary care physicians who offer risk profiling. J Gen Intern Med 2011 Aug;26(8):834-840 [FREE Full text] [doi: 10.1007/s11606-011-1651-7] [Medline: 21311998]
101. Hock KT, Christensen KD, Yashar BM, Roberts JS, Gollust SE, Uhlmann WR. Direct-to-consumer genetic testing: an assessment of genetic counselors' knowledge and beliefs. Genet Med 2011 Apr;13(4):325-332 [FREE Full text] [doi: 10.1097/GIM.0b013e3182011636] [Medline: 21233722]
102. Mainous A, Johnson S, Chirina S, Baker R. Academic family physicians' perception of genetic testing and integration into practice: a CERA study. Fam Med 2013 Apr;45(4):257-262. [Medline: 23553089]
103. Noss R, Mills R, Callanan N. The incorporation of predictive genomic testing into genetic counseling programs. J Genet Couns 2014 Aug;23(4):671-678. [doi: 10.1007/s10897-014-9699-2] [Medline: 24584888]
104. Ormond KE, Hudgins L, Ladd JM, Magnus DM, Greely HT, Cho MK. Medical and graduate students' attitudes toward personal genomics. Genet Med 2011 May;13(5):400-408. [doi: 10.1097/GIM.0b013e31820562f6] [Medline: 21270640]
105. Powell KP, Cogswell WA, Christianson CA, Dave G, Verma A, Eubanks S, et al. Primary care physicians' awareness, experience and opinions of direct-to-consumer genetic testing. J Genet Couns 2012 Feb;21(1):113-126. [doi: 10.1007/s10897-011-9390-9] [Medline: 21769569]
106. Powell KP, Christianson CA, Cogswell WA, Dave G, Verma A, Eubanks S, et al. Educational needs of primary care physicians regarding direct-to-consumer genetic testing. J Genet Couns 2012 Jun;21(3):469-478. [doi: 10.1007/s10897-011-9471-9] [Medline: 22207397]
107. Salm M, Abbate K, Appelbaum P, Ottman R, Chung W, Marder K, et al. Use of genetic tests among neurologists and psychiatrists: knowledge, attitudes, behaviors, and needs for training. J Genet Couns 2014 Apr;23(2):156-163 [FREE Full text] [doi: 10.1007/s10897-013-9624-0] [Medline: 23793969]
108. Brierley KL, Bonadies DC, Moyer A, Matloff ET. "Would you test your children without their consent?" and other sticky dilemmas in the field of cancer genetic testing. Fam Cancer 2014 Sep;13(3):345-350. [doi: 10.1007/s10689-014-9723-6] [Medline: 24804937]
109. Brett GR, Metcalfe SA, Amor DJ, Halliday JL. An exploration of genetic health professionals' experience with direct-to-consumer genetic testing in their clinical practice. Eur J Hum Genet 2012 Aug;20(8):825-830 [FREE Full text] [doi: 10.1038/ejhg.2012.13] [Medline: 22317975]
110. Howard HC, Borry P. Survey of European clinical geneticists on awareness, experiences and attitudes towards direct-to-consumer genetic testing. Genome Med 2013;5(5):45 [FREE Full text] [doi: 10.1186/gm449] [Medline: 23697740]
111. Ohata T, Tsuchiya A, Watanabe M, Sumida T, Takada F. Physicians' opinion for 'new' genetic testing in Japan. J Hum Genet 2009 Apr;54(4):203-208. [doi: 10.1038/jhg.2009.11] [Medline: 19300458]
112. Ram S, Russell B, Gubb M, Taylor R, Butler C, Khan I, et al. General practitioner attitudes to direct-to-consumer genetic testing in New Zealand. N Z Med J 2012 Oct 26;125(1364):14-26. [Medline: 23242394]
113. Arribas-Ayllon M, Sarangi S, Clarke A. Promissory accounts of personalisation in the commercialisation of genomic knowledge. Commun Med 2011;8(1):53-66. [Medline: 22616356]
114. Hennen L, Sauter A, Van Den Cruyce E. Direct to consumer genetic testing: Insights from an internet scan. New Genet Soc 2010 Jun 21;29(2):167-186. [doi: 10.1080/14636778.2010.484232]
115. Goddard KA, Robitaille J, Dowling NF, Parrado AR, Fishman J, Bradley LA, et al. Health-related direct-to-consumer genetic tests: a public health assessment and analysis of practices related to Internet-based tests for risk of thrombosis. Public Health Genomics 2009;12(2):92-104 [FREE Full text] [doi: 10.1159/000176794] [Medline: 19039253]
116. Liu Y, Pearson YE. Direct-to-Consumer Marketing of Predictive Medical Genetic Tests: Assessment of Current Practices and Policy Recommendations. J Public Policy Mark 2008;27(2):131-148.
117. Sterling R. The on-line promotion and sale of nutrigenomic services. Genet Med 2008 Nov;10(11):784-796. [doi: 10.1097/GIM.0b013e31818c0441] [Medline: 18941418]
118. Borry P, Howard HC, Sénécal K, Avard D. Health-related direct-to-consumer genetic testing: a review of companies' policies with regard to genetic testing in minors. Fam Cancer 2010 Mar;9(1):51-59. [doi: 10.1007/s10689-009-9253-9] [Medline: 19488835]
119. Howard HC, Avard D, Borry P. Are the kids really all right? Direct-to-consumer genetic testing in children: are company policies clashing with professional norms? Eur J Hum Genet 2011 Nov;19(11):1122-1126 [FREE Full text] [doi: 10.1038/ejhg.2011.94] [Medline: 21629297]
120. Kechagia S, Mai Y, Vidalis T, Patrinos GP, Vayena E. Personal genomics in Greece: an overview of available direct-to-consumer genomic services and the relevant legal framework. Public Health Genomics 2014;17(5-6):299-305. [doi: 10.1159/000366175] [Medline: 25278084]

121. Lachance CR, Erby LAH, Ford BM, Allen VC, Kaphingst KA. Informational content, literacy demands, and usability of websites offering health-related genetic tests directly to consumers. Genet Med 2010 May;12(5):304-312 [FREE Full text] [doi: 10.1097/GIM.0b013e3181dbd8b2] [Medline: 20386454]
122. Lewis NP, Treise D, Hsu SI, Allen WL, Kang H. DTC genetic testing companies fail transparency prescriptions. New Genetics and Society 2011 Dec;30(4):291-307. [doi: 10.1080/14636778.2011.600434]
123. Geransar R, Einsiedel E. Evaluating online direct-to-consumer marketing of genetic tests: informed choices or buyers beware? Genet Test 2008 Mar;12(1):13-23.
124. Covolo L, Rubinelli S, Orizio G, Gelatti U. Misuse (and abuse?) of the Concept of Empowerment.The Case of Online Offer of Predictive Direct-to-consumer Genetic Tests. J Public Health Res 2012 Feb 17;1(1):7-10 [FREE Full text] [doi: 10.4081/jphr.2012.e3] [Medline: 25170440]
125. Gollust SE, Wilfond BS, Hull SC. Direct-to-consumer sales of genetic services on the Internet. Genet Med 2003;5(4):332-337. [doi: 10.1097/01.GIM.0000076972.83711.48] [Medline: 12865763]
126. Harris A, Kelly SE, Wyatt S. Counseling customers: emerging roles for genetic counselors in the direct-to-consumer genetic testing market. J Genet Couns 2013 Apr;22(2):277-288. [doi: 10.1007/s10897-012-9548-0] [Medline: 23093333]
127. Singleton A, Erby LH, Foisie KV, Kaphingst KA. Informed choice in direct-to-consumer genetic testing (DTCGT) websites: a content analysis of benefits, risks, and limitations. J Genet Couns 2012 Jun;21(3):433-439 [FREE Full text] [doi: 10.1007/s10897-011-9474-6] [Medline: 22194036]
128. Einsiedel EF, Geransar R. Framing genetic risk: trust and credibility markers in online direct-to-consumer advertising for genetic testing. New Genetics and Society 2009 Dec;28(4):339-362. [doi: 10.1080/14636770903314509]
129. Kutz G. United States Government Accountability Office. 2006. Nutrigenetic Testing:Tests Purchased from Four Web Sites Mislead Consumers URL: http://www.gao.gov/new.items/d06977t.pdf [accessed 2015-02-12] [WebCite Cache ID 6WI1Hs3Tg]
130. Adams SD, Evans JP, Aylsworth AS. Direct-to-consumer genomic testing offers little clinical utility but appears to cause minimal harm. N C Med J 2013;74(6):494-498 [FREE Full text] [Medline: 24316774]
131. Bloss CS, Topol EJ, Schork NJ. Association of direct-to-consumer genome-wide disease risk estimates and self-reported disease. Genet Epidemiol 2012 Jan;36(1):66-70 [FREE Full text] [doi: 10.1002/gepi.20664] [Medline: 22127769]
132. Buitendijk GHS, Amin N, Hofman A, van Duijn CM, Vingerling JR, Klaver CCW. Direct-to-consumer personal genome testing for age-related macular degeneration. Invest Ophthalmol Vis Sci 2014 Oct;55(10):6167-6174. [doi: 10.1167/iovs.14-15142] [Medline: 25146986]
133. Imai K, Kricka LJ, Fortina P. Concordance study of 3 direct-to-consumer genetic-testing services. Clin Chem 2011 Mar;57(3):518-521 [FREE Full text] [doi: 10.1373/clinchem.2010.158220] [Medline: 21159896]
134. Johnson AD, Bhimavarapu A, Benjamin EJ, Fox C, Levy D, Jarvik GP, et al. CLIA-tested genetic variants on commercial SNP arrays: potential for incidental findings in genome-wide association studies. Genet Med 2010 Jun;12(6):355-363 [FREE Full text] [doi: 10.1097/GIM.0b013e3181e1e2a9] [Medline: 20556870]
135. Kalf RRJ, Mihaescu R, Kundu S, de KP, Green RC, Janssens CJW. Variations in predicted risks in personal genome testing for common complex diseases. Genet Med 2014 Jan;16(1):85-91 [FREE Full text] [doi: 10.1038/gim.2013.80] [Medline: 23807614]
136. Kido T, Kawashima M, Nishino S, Swan M, Kamatani N, Butte AJ. Systematic evaluation of personal genome services for Japanese individuals. J Hum Genet 2013 Nov;58(11):734-741. [doi: 10.1038/jhg.2013.96] [Medline: 24067293]
137. Mihaescu R, van HM, Sijbrands EJG, Uitterlinden AG, Witteman JCM, Hofman A, et al. Evaluation of risk prediction updates from commercial genome-wide scans. Genet Med 2009 Aug;11(8):588-594. [doi: 10.1097/GIM.0b013e3181b13a4f] [Medline: 19636253]
138. Ng PC, Murray SS, Levy S, Venter JC. An agenda for personalized medicine. Nature 2009 Oct 8;461(7265):724-746. [doi: 10.1038/461724a] [Medline: 19812653]
139. Palomaki GE, Melillo S, Marrone M, Douglas MP. Use of genomic panels to determine risk of developing type 2 diabetes in the general population: a targeted evidence-based review. Genet Med 2013 Aug;15(8):600-611. [doi: 10.1038/gim.2013.8] [Medline: 23492876]
140. Swan M. Multigenic condition risk assessment in direct-to-consumer genomic services. Genet Med 2010 May;12(5):279-288. [doi: 10.1097/GIM.0b013e3181d5f73b] [Medline: 20474084]
141. Orizio G, Merla A, Schulz PJ, Gelatti U. Quality of online pharmacies and websites selling prescription drugs: a systematic review. J Med Internet Res 2011;13(3):e74 [FREE Full text] [doi: 10.2196/jmir.1795] [Medline: 21965220]
142. Covolo L, Capelli M, Ceretti E, Feretti D, Caimi L, Gelatti U. Nutritional supplements for diabetes sold on the internet: business or health promotion? BMC Public Health 2013 Aug 26;13:777.
143. Human Genetics Commission. 2010. A common framework of principles for direct-to-consumer genetic testing services URL: http://osp.od.nih.gov/sites/default/files/resources/Human_Genetics_Commission_0.pdf [WebCite Cache ID 6WI1Mhpal]
144. Howard HC, Borry P. Is there a doctor in the house? : The presence of physicians in the direct-to-consumer genetic testing context. J Community Genet 2012 Apr;3(2):105-112 [FREE Full text] [doi: 10.1007/s12687-011-0062-0] [Medline: 22109907]

145. Draugalis JR, Coons SJ, Plaza CM. Best practices for survey research reports: a synopsis for authors and reviewers. Am J Pharm Educ 2008 Feb 15;72(1):11 [FREE Full text] [Medline: 18322573]
146. Janssens CJ, van Duijn CM. An epidemiological perspective on the future of direct-to-consumer personal genome testing. Investig Genet 2010;1(1):10 [FREE Full text] [doi: 10.1186/2041-2223-1-10] [Medline: 21092344]
147. Global Industry Analysts, Inc. 2013. Direct-to-Consumer (DTC) Genetic Testing: A Global Strategic Business Report URL: http://www.prweb.com/releases/DTC_genetic_testing/direct_to_consumer_tests/prweb9780295.htm [accessed 2015-02-12] [WebCite Cache ID 6WI1RVqFp]

Abbreviations

BRCA: breast cancer susceptibility protein
DTC-GT: direct-to-consumer genetic testing
FDA: Food and Drug Administration
GAO: Government Accountability Office
GT: genetic testing/test
SNP: single nucleotide polymorphism

Edited by G Eysenbach; submitted 10.03.15; peer-reviewed by Y Mai, J Jiang; comments to author 05.08.15; revised version received 12.10.15; accepted 16.10.15; published 14.12.15.

Please cite as:
Covolo L, Rubinelli S, Ceretti E, Gelatti U
Internet-Based Direct-to-Consumer Genetic Testing: A Systematic Review
J Med Internet Res 2015;17(12):e279
URL: http://www.jmir.org/2015/12/e279/
doi:10.2196/jmir.4378
PMID:26677835

Original Paper

Barriers and Facilitators to Online Portal Use Among Patients and Caregivers in a Safety Net Health Care System: A Qualitative Study

Lina Tieu[1], MPH; Urmimala Sarkar[1], MD, MPH; Dean Schillinger[1], MD; James D Ralston[2], MD, MPH; Neda Ratanawongsa[1], MD, MHS; Rena Pasick[3], DrPH; Courtney R Lyles[1], PhD

[1]Division of General Internal Medicine, University of California, San Francisco, San Francisco, CA, United States

[2]Group Health Research Institute, Seattle, WA, United States

[3]Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Lina Tieu, MPH
Division of General Internal Medicine
University of California, San Francisco
DGIM at SFGH, Box 1364
1001 Potrero Ave, Bldg 10, Ward 13
San Francisco, CA, 94143
United States
Phone: 1 415 206 7878
Fax: 1 415 206 7880
Email: Lina.Tieu@ucsf.edu

Abstract

Background: Patient portals have the potential to support self-management for chronic diseases and improve health outcomes. With the rapid rise in adoption of patient portals spurred by meaningful use incentives among safety net health systems (a health system or hospital providing a significant level of care to low-income, uninsured, and vulnerable populations), it is important to understand the readiness and willingness of patients and caregivers in safety net settings to access their personal health records online.

Objective: To explore patient and caregiver perspectives on online patient portal use before its implementation at San Francisco General Hospital, a safety net hospital.

Methods: We conducted 16 in-depth interviews with chronic disease patients and caregivers who expressed interest in using the Internet to manage their health. Discussions focused on health care experiences, technology use, and interest in using an online portal to manage health tasks. We used open coding to categorize all the barriers and facilitators to portal use, followed by a second round of coding that compared the categories to previously published findings. In secondary analyses, we also examined specific barriers among 2 subgroups: those with limited health literacy and caregivers.

Results: We interviewed 11 patients and 5 caregivers. Patients were predominantly male (82%, 9/11) and African American (45%, 5/11). All patients had been diagnosed with diabetes and the majority had limited health literacy (73%, 8/11). The majority of caregivers were female (80%, 4/5), African American (60%, 3/5), caregivers of individuals with diabetes (60%, 3/5), and had adequate health literacy (60%, 3/5). A total of 88% (14/16) of participants reported interest in using the portal after viewing a prototype. Major perceived barriers included security concerns, lack of technical skills/interest, and preference for in-person communication. Facilitators to portal use included convenience, health monitoring, and improvements in patient-provider communication. Participants with limited health literacy discussed more fundamental barriers to portal use, including challenges with reading and typing, personal experience with online security breaches/viruses, and distrust of potential security measures. Caregivers expressed high interest in portal use to support their roles in interpreting health information, advocating for quality care, and managing health behaviors and medical care.

Conclusions: Despite concerns about security, difficulty understanding medical information, and satisfaction with current communication processes, respondents generally expressed enthusiasm about portal use. Our findings suggest a strong need for training and support to assist vulnerable patients with portal registration and use, particularly those with limited health literacy.

Efforts to encourage portal use among vulnerable patients should directly address health literacy and security/privacy issues and support access for caregivers.

(J Med Internet Res 2015;17(12):e275) doi:10.2196/jmir.4847

KEYWORDS

personal health records; electronic health records; chronic disease; caregivers; health literacy; safety-net providers

Introduction

Over the past few decades, there has been a surge in the use of electronic health records (EHRs) in the United States, spurred by the Health Information Technology for Economic and Clinical Health (HITECH) Act [1] “Meaningful Use” financial incentive program [2]. As Meaningful Use has moved into its fourth year, its incentives have promoted the rapid uptake of online patient portals by health systems nationwide [3], allowing patients to access laboratory test results, view visit summaries, and email their providers. Patient portals have been touted as a way to support self-management for chronic diseases by promoting disease awareness and knowledge, self-efficacy, and improvements in health behaviors and communication [4-9]. Early evidence also linked portal use to better outcomes such as risk factor control for diabetes [10,11].

Despite the potential of portals to promote patient engagement and improve self-management [12], there is evidence that not all patient subgroups use portals similarly. Among integrated health care systems with well-established portals, there is consistent evidence that racial/ethnic minorities and patients with lower income, education, and health literacy are significantly less likely to use available portal websites [10,13-19]. Although some studies tried to elucidate general barriers to portal use, such as computer/Internet access, attitudes/preferences, awareness, and security/privacy of information [5,7,13,20], few studies to date have reported on health literacy as a barrier to portal use and interpretation [14,21] and none have reported specific barriers among individuals with limited health literacy. Moreover, caregivers—although recognized as increasingly important in the care of those with chronic illness [22]—have often been left out of studies examining portal use to date.

A safety net hospital or health system provides a significant level of care to low-income, uninsured, and vulnerable populations. Although there have been early adopters in the field [23,24], the use of online patient portals is new territory for many safety net health systems because many have just completed implementation of their EHRs. In 2012, 40% of community clinics and health centers in California reported at least a basic EHR system [25] and even fewer had provided patients access to their personal health record information online. Given the potential for portals to improve self-management, it is important to assess the readiness and willingness of patients and caregivers to access personal health records online in safety net settings—especially because patient interest is high [26,27].

In this qualitative study, we sought to elucidate the barriers and facilitators to use of a patient portal in anticipation of portal implementation in an urban, safety net primary care clinic.

Methods

Research Setting

The study was conducted at San Francisco General Hospital (SFGH), a safety net hospital in the San Francisco Health Network system. From December 2013 to September 2014, we recruited participants to gain a pool of individuals with a wide range of health knowledge and engagement. Recruitment sites included (1) the General Medicine Clinic (GMC), a primary care clinic serving more than 6500 patients, most of whom are uninsured (32%) or on Medicaid (39%); (2) a diabetes support group led by diabetes nurses; and (3) a diabetes education class that GMC patients were referred to at the hospital. GMC began exclusively using an EHR in June of 2013. At the time of the interviews, the SFGH-wide patient portal was not yet launched; rollout was scheduled for early 2015.

Sampling Procedure

We recruited patients through an electronic query of patients with upcoming clinic or diabetes group appointments. To recruit participants identified as having upcoming clinic appointments, study staff approached potential participants before or after their appointments, explained the study, and recruited interested participants. To recruit participants identified as being enrolled in group sessions, study staff attended the sessions, described the study to the group, and recruited interested participants. Caregivers, defined as someone playing a role in the management of a patient’s health other than the patient or the medical provider, were recruited by provider referral of someone who attended medical visits with or communicated with a provider on behalf of a patient. Participants were eligible for the study if they were (1) English-speaking, (2) not cognitively impaired, and (3) diagnosed with a chronic disease or the caregiver of such a patient. We focused on patients with chronic illnesses because portal use may be particularly useful in supporting ongoing self-management. We included only participants who expressed some interest in using the Internet overall to manage their health care, unless accompanied by a caregiver who expressed such interest, because we felt that this represented a realistic sample of individuals who would be potentially interested in and able to use the portal when it launched.

Data Collection Procedure

During recruitment, we administered a short questionnaire to gather information on demographics (age, race/ethnicity, gender); diagnosis of a chronic disease (heart disease, diabetes, high blood pressure, heart failure, asthma/chronic obstructive pulmonary disease [COPD], and/or chronic kidney disease); interest in using the Internet to manage health care at SFGH (high, some, none, or don’t know/need more information); and frequency of current Internet use (daily, weekly, monthly or less, or none). Finally, we administered a previously validated one-item health literacy scale regarding how confident participants were filling out medical forms on their own (not at all, a little bit, somewhat, quite a bit, extremely) [28] because this has been shown to be predictive of portal use in our previous quantitative work [14]. We classified participants noting any lack of confidence in filling out forms as having limited health literacy.

We conducted semi-structured in-depth interviews with 11 patients and 5 caregivers, two of which were dyad interviews with both a patient and his or her respective caregiver (Multimedia Appendix 1, Interview Guide). Although the interviews included discussion of current health status, health behaviors, and health care utilization, emphasis was placed on prior use of the Internet and specific interest in the use of a patient portal website for health management, informed by the Theory of Acceptance and Use of Technology [29,30]. To provide a visual example of a patient portal interface, participants were shown screenshots of a sample patient portal interface on paper, including the log-in, test results, and visit summary features. Participants were asked to state whether they thought they would use the portal website in their own care and what features of the portal website were of most interest to them. Caregivers were asked to discuss the potential impact proxy access to a patient's personal health records would have on their role. The interviews were transcribed and deidentified before analysis.

Data Analysis

Using data from the questionnaires, we summarized the participant demographics and Internet and portal use responses.

Authors CL and LT read the interview transcripts in their entirety before independently analyzing them. We used an interpretive description approach [31,32] to analyze the transcripts, using inductive and deductive coding techniques. Coding was done using Atlas.ti 7 software [33]. First, we used inductive open coding to identify all emerging themes and subthemes that participants mentioned during the discussions [34,35]. To assure quality of the analysis and to uphold the constant comparison open coding approach, CL and LT met regularly to discuss the thematic findings. When there was disagreement, US established agreement on codes. The entire team reviewed and provided comments on the final codebook.

A main goal in this study was to determine if the barriers and facilitators in our safety net setting were similar or different from previously published literature on portal use in other settings—the majority of which have been conducted in integrated health care settings. For this reason, we recoded all the transcripts in a deductive manner to be able to determine if the categories and severity of barriers and facilitators that emerged from our analysis were comparable to previously published work [5,13]. Study staff halted further enrollment in the study after a consensus that thematic saturation had been reached.

Because there was a clear indication of clustering of themes by health literacy status, the transcripts were re-examined by self-reported health literacy status in a secondary exploratory analysis. Although we did not purposively sample for limited health literacy status, we were able to generate some hypotheses for additional types of barriers for patients in this group. We also examined the caregiver transcripts independently from the patient interviews in a similar exploratory fashion.

Finally, we summarized basic usability and accessibility comments as participants looked at screenshots of a hypothetical portal website to understand how they might use the website in the future. This included identifying patients reporting interest in using specific portal features.

Results

Enrollment

A total of 45 individuals were approached about the study. Of those approached, 25 (56%) expressed interest in the study, 7 (16%) declined citing lack of interest in computer use, and 13 (29%) declined due to unknown or other reasons (too busy, uninterested in research). Of the 25 who expressed initial interest, 3 subsequently could not be reached, 5 stated they were too busy to schedule an interview, and 1 declined an interview due to emerging health issues. We enrolled 16 participants in the study before reaching thematic saturation.

Description of Sample

Participants in the study were predominately male (10/16, 63%) and ethnically diverse (50%, 8/16 African American; 19%, 3/16 Latino; 19%, 3/16 Asian or Pacific Islander; 13%, 2/16 white). All patients in the sample were diagnosed with diabetes, 60% (3/5) of caregivers cared for individuals with diabetes and 20% (1/5) of caregivers cared for individuals with multiple chronic conditions, including hypertension, heart disease, chronic kidney disease, and COPD. The mean age of the sample was 56 years (SD 11). More than half (10/16, 63%) of participants had limited health literacy. Overall, participants reported high experience and interest in Internet use: 56% (9/16) expressed high interest in using the Internet to manage their health care and 69% (11/16) were daily users of the Internet. All but 2 participants reported at least occasional Internet use (Table 1).

Table 1. Participant demographics.

Characteristic	Overall n=16	Patients n=11	Caregivers n=5
Age (years), mean (SD)	56 (11)	57 (8)	52 (16)
Gender, n (%)			
Male	10 (63)	9 (82)	1 (20)
Female	6 (38)	2 (18)	4 (80)
Race/Ethnicity, n (%)			
Black or African American	8 (50)	5 (45)	3 (60)
Hispanic/Latino	3 (19)	3 (27)	0 (0)
Asian or Pacific Islander	3 (19)	2 (18)	1 (20)
White or Caucasian	2 (13)	1 (9)	1 (20)
Role, n (%)			
Patient	11 (69)	11 (100)	N/A
Caregiver	5 (31)	N/A	5 (100)
Health literacy status, n (%)			
Limited	10 (63)	8 (73)	2 (40)
Adequate	6 (38)	3 (27)	3 (60)
Interest in using Internet to manage health, n (%)			
High	9 (56)	5 (45)	4 (80)
Some	4 (25)	3 (27)	1 (20)
None	3 (19)	2 (18)	0 (0)
Don’t know	1 (6)	1 (9)	0 (0)
Frequency of Internet use, n (%)			
Daily	11 (69)	7 (64)	4 (80)
Weekly	2 (13)	1 (9)	1 (20)
Every 2-3 Weeks	1 (6)	1 (9)	0 (0)
Never	2 (13)	2 (18)	0 (0)
Internet access, n (%)			
Personal computer	13 (81)	9 (82)	4 (80)
Personal mobile phone	10 (63)	5 (45)	5 (100)
Computer in public setting	2 (13)	2 (18)	0 (0)

Major Categories

Overall, the 5 major categories characterizing the barriers and facilitators for portal use were similar to the previously published research on this topic: (1) computer or Internet access, (2) technological skills and interest, (3) security and privacy of information, (4) patient-provider relationship, and (5) chronic illness self-management.

The secondary analysis by health literacy status showed a much higher prevalence of barriers for participants with limited health literacy within several of these categories. These findings are highlighted subsequently.

Computer or Internet Access

Overall, the majority of participants reported having consistent and easy access to a personal computer, tablet, or phone. All participants but one owned a computer or mobile device with Internet access. Two participants accessed the Internet solely using a mobile device or tablet, whereas 2 participants with limited health literacy reported accessing a computer or the Internet in public areas, such as a library, classroom, or through a friend:

> *I go to the library sometimes or a friend’s house or something there, or when I get with the tutor or something and they’re teaching me something, they’ll teach me on their computer or stuff like that.* [Male

> patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

In addition, 2 individuals with limited health literacy expressed concerns about the affordability of the Internet, particularly concerning the cost of mobile data:

> *I don't have a camera phone. Plus, I don't have $35.00 that happens to be the monthly fee.* [Male patient, age 51-55 years, Asian or Pacific Islander, with limited health literacy]

Technological Skills and Interest

Although we required some interest in technology in order to enter the study, there was a wide spectrum of technological proficiency, from very limited experience using computers to formal schooling in computer-related fields. Although 3 participants mentioned age as a limiting factor in being able to keep up with the "new generation" of technology use, overall, participants reported using the computer and Internet for a variety of tasks, including communication with friends and family, research, banking, and shopping. Although the majority of participants with adequate health literacy had advanced knowledge of computers, all participants with limited health literacy described their skill levels as being potentially limiting with respect to using a portal:

> *[My doctor] knows that I'm into computers. I'm a major in computers so [using the portal] is up my alley.* [Male patient, age 41-45 years, Asian or Pacific Islander, with adequate health literacy]
>
> *Yes, [sending an email] would teach me how to type and all that.* [Patient, age 46-50 years, African American, with limited health literacy]

Although some difficulties reflected issues of cognitive overload, such as difficulty remembering passwords, others represented a lack of basic computer skills. Five participants reported comfort using passwords, but expressed that it was often difficult to remember them, exacerbated by the requirement of websites to change passwords at specific intervals. One participant with limited health literacy also reported difficulty creating passwords, particularly understanding the requirements that websites impose to promote password security, and expressed confusion about the requirement to create your own password:

> *That's another thing because you got to have so many words and letters. You know, characters, so how do you distinguish that? I mean you say characters, are they letters?...Where do you get that at? Where do you get the password at?* [Male caregiver, age 56-60 years, African American, with limited health literacy]

Another participant expressed hesitancy using any websites that require a username and password, preferring to complete transactions such as shopping and banking in-person or over the phone:

> *Usually when I get to those, I don't log in...it won't let me in, I won't get on it.* [Male patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

Security and Privacy of Information

Most participants expressed concern about their health information being online, although there was nuanced understanding of both the benefits and risks of accessing information online. More specifically, 7 participants noted the vulnerability of online systems to hackers:

> *You hear so many instances where information has been compromised. I mean, the military can be compromised.* [Male patient, age 56-60 years, African American, with limited health literacy]

Participants were also concerned about the confidentiality of their health information, particularly sensitive diagnoses and medications. At the same time, 4 participants were unconcerned about security breaches. Participants believed that their personal information was already publicly available through online searches. In addition, they expressed that hackers would find little value in their personal information because they felt unimportant or lacked employment for which the leak of sensitive information would be a threat. Despite concerns about security, 2 participants noted a trust in the ability of a complicated password to improve the security of the hypothetical portal:

> *I guess for me, more secure is to give a special password. One key. One key to keep it confidential to go in.* [Male patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

Two participants with limited health literacy described past experiences with computer viruses or information breaches, contributing to their current concerns about online security:

> *Hackers getting [into] everything...I had to change banks because...they had everything—my name and address—my mom's maiden name.* [Male patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

In addition, they described their distrust of potential security measures, including the ability of researchers and industry members to access their health information:

> *Regardless of what a person says that this site is secured and all that, I just don't believe it...It's not only hospitals but pharmaceutical and every researcher will tap into my information. That's the thing that I worry about.* [Male patient, age 51-55 years, Asian or Pacific Islander, with limited health literacy]

Patient-Provider Relationship

All but one participant noted the benefits of portal use, mainly the option to securely message their provider to get answers to questions not requiring a visit. Five participants also discussed how accessing their personal health records would improve the effectiveness of their in-person visits. Participants with adequate health literacy possessed a more advanced understanding about how improving transparency and knowledge about their health could improve visits by allowing them to ask their providers more specific questions about their diet, exercise, medication, and other management topics:

> *When you go to a visit, you can ask more specific questions. You can say, "I looked at my labs, and I saw that my A1C was dah, dah, dah." I know that that represents the fact that I have not been as good with my management as I should be.* [Male patient, age 61-65 years, African American, with adequate health literacy]

However, 4 participants emphasized the value of in-person communication and did not want online secure messaging to diminish or interfere with ongoing in-person communication with their providers. They expressed concern over technology replacing their health care providers:

> *I wouldn't want anything like...I've seen those things on television where they got the doctor and you see the doctor on the computer screen and stuff. Is that kind of like what you're talking about?* [Male patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

Some participants believed providers would have time to interact with them via the portal, especially through secure messaging. On the other hand, 6 participants, including 5 with limited health literacy, were more skeptical, noting past instances in which they had tried to email their providers to no avail:

> *Everything is now computer, so yes they would have time. When they're sitting right there looking at your chart on the computer, that's their time they email about the patient.* [Female caregiver, age 51-55 years, African American, with adequate health literacy]
>
> *Well, unless I'm missing something very basic, email just doesn't seem to work...it's like it goes into this pot.* [Male patient, age 61-65 years, African American, with limited health literacy]

Chronic Illness Self-Management

All participants expressed positive statements about the portal in relation to their health management, including coordination of care and health promotion. For diabetes patients, the option to view past test results was noted as being particularly useful in tracking progress and adjusting health behaviors, such as diet and exercise:

> *Particularly things like the lab would be good...because then you could not fool yourself. You would have your regular [glucometer] record and your quarterly A1C record to compare so you could see the connection and also give the physician the same ability.* [Male patient, age 61-65 years, African American, with adequate health literacy]

Ten participants noted that using a portal would save them time in managing their health. In addition, patients felt that using the portal would promote patient-driven communication by improving the ability to seek medical advice in between visits on topics including medication side effects, test results, symptoms, and new treatments seen in the media.

> *[If] I had a consultation with my pharmacist and they're telling me of the side effects to watch out with some medications I'm taking...[and] I have one of those side effects, I might discuss it with a doctor on email. That would be really helpful.* [Female patient, age 46-50 years, white, with limited health literacy]

Three participants with limited health literacy were particularly enthusiastic about the option to check their future appointments online, noting past instances where they had missed appointments because they had forgotten or hadn't received the proper notification:

> *Because sometimes they don't get [the appointment reminder] out on time, so they end up at the last minute sending it out or something, and then [you] find out you had an appointment.* [Male patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

Strong Interest Among Caregivers

Of the 5 caregivers in our sample, 4 were female; 2 were romantic partners, 2 were children of a patient, and 1 was an in-home professional caregiver. All major themes were similar among caregivers and patients, but interest in using the portal seemed stronger among the few caregivers in our sample, primarily because they were already highly engaged in health care management tasks. In the dyad interviews with patients and caregivers, there was no hesitancy among patients to share their personal health records with their caregivers. It was clear that caregivers in this study already played a strong role in communicating with providers and portals would be a logical extension of their role. The majority of caregivers that we interviewed were already informally in contact with a patient's provider via individual email accounts (ie, not through a portal website) and spoke of their experience positively. Caregivers also noted the importance of their role as interpreters of health information:

> *I think he would be looking at [the portal] with me but he just doesn't understand so I would just have to relay the message.* [Female caregiver for parent, age 21-25 years, Asian or Pacific Islander, with adequate health literacy]

In addition, caregivers described their role as advocates in the care of patients, particularly in advocating that patients not miss out on critical in-person visits as a result of the portal:

> *I would like [Patient] to go have his visit with the doctor and he loves coming to the doctor and seeing his doctors. Do not take that away.* [Female in-home supportive service caregiver, age 51-55 years, African American, with adequate health literacy]

Caregivers also discussed the potential for portal use to improve their ability to monitor and promote improvements in health behavior.

> *To be able to monitor him even better 'cause now I can go on there, I can look, I can see the results, show him what it's saying in case he forgets, and let him know, this is what you should do. You need to stop doing this and do this.* [Female caregiver for parent, age 56-60 years, African American, with adequate health literacy]

Caregivers discussed using the patient portal both independently and in tandem with patients through the patient's account. As opposed to creating a separate proxy account, one caregiver noted that she would create an account for her parent, which she and a sibling would both use to access the portal:

> *I'm probably going to make my dad one [portal account] and stuff like that...I'll teach my brother.* [Female caregiver for parent, age 21-25 years, Asian or Pacific Islander, with adequate health literacy]

Portal Usability

After viewing hypothetical portal screenshots, participants expressed some challenges with the medical terminology and lack of language-appropriate information, but thought the portal layout was otherwise straightforward and comprehensive (Textbox 1).

Textbox 1. Perceptions of hypothetical portal example.

Weaknesses

1. Difficulty understanding portal content

> *Yes, and then the lab result, even though I won't understand most of it.* [Male patient, age 41-45 years, Asian or Pacific Islander, with adequate health literacy]
>
> *Probably to see a blood test result. I wouldn't really—unless somebody explained it, I wouldn't know what I was looking at, really. It's like diagnosing your car, tells you all this stuff but then you don't know what it is. I got so much stuff.* [Male patient, age 56-60 years, Hispanic/Latino, with limited health literacy]

2. Language access or limited English proficiency

> *Is there any other options like other languages that you can kind of change the message to? Like not permanently but let's just say that day, if I teach my dad how to go online and he can look up for himself, like that day when he go on, can he click a certain button that's not that hard for him to change it, let's say to Vietnamese.* [Female caregiver for parent, age 21-25 years, Asian or Pacific Islander, with adequate health literacy]
>
> *To be honest with you, unless it's something interesting to go into that health thing, then I would go. For example, if that's in my language, I would go.* [Male patient, age 51-55 years, Asian or Pacific Islander, with limited health literacy]

Strengths

1. Hypothetical portal simple and clear

> *Yes. It's much easier than when I was in school. That portal was awful.* [Male patient, age 41-45 years, Asian or Pacific Islander, with adequate health literacy]
>
> *Well it seems really self-explanatory. It's like really basic, just all right there. I can't think of anything to add to it.* [Female patient, age 46-50 years, white, with limited health literacy]

Interest in Portal Use

After seeing the example screenshots of the future patient portal, 88% (14/16) of participants reported a willingness to use the future portal website to manage their health care. Looking at specific features, there was highest interest in accessing laboratory results (81%, 13/16), appointments (81%, 13/16), and visit summaries (81%, 13/16).

Discussion

Principal Findings

Among a diverse group of patients and caregivers in a safety net clinic, we identified significant barriers to portal use, including concerns about security and privacy, limited technological proficiency, and a desire to preserve in-person aspects of existing patient-provider relationships—most often among patients with limited health literacy. Recruiting only those who expressed at least some interest in using the Internet for health management, our findings are likely conservative in that they represent some of the more engaged patients within our safety net health care system. The majority of participants in our study were African American and male, characteristics which have both been associated with lower portal use [13,14,17,26,36]. Nevertheless, it is important to note there was overall enthusiasm among these participants about the potential of a patient portal to improve aspects of health monitoring, patient-provider relationships, and caregiver burden. This is consistent with interest [21,26,27] and benefits described among low-income patients in past studies [5,21,27].

The overall categories of barriers to portal use in our study were consistent with previous studies: concerns about security [5,37], difficulty understanding medical information [5,21], the desire to preserve verbal communication and in-person contact [5,20,38], and the burden of portal use on clinician workloads [5]. However, because our sample included predominantly patients with limited health literacy, our findings uncovered more pronounced aspects of these barriers in safety net settings, such as access to computers in public settings; negative past experiences with technology, including security breaches and viruses; and a lack of more rudimentary computer skills. In particular, a distrust of potential security measures to prevent access of personal health information by hackers, researchers, and others unauthorized by the patient may hinder patient portal use among safety net patients. This is consistent with past studies indicating that individuals with limited health literacy are less likely to sign onto a portal website [14,39], use portal messaging functions [39], identify blood sugar test results as out of range [19], and contact a provider to discuss abnormal test results from an online portal [19]. Public computer use coupled with relatively widespread security concerns may be particularly relevant to address among patients with lower socioeconomic status, especially because previous studies have suggested that only a minority of a general patient population express hesitancy to use portals because of security [37,40]. Past research has shown that older, low-income patients desire assistance in interpreting their personal health information [41]. Our findings show that health literacy is a major barrier among younger populations within the safety net as well.

Although our sample of caregivers was small (limiting our ability to make strong inference), our findings also imply that there is a potential supportive role for caregivers to facilitate portal use in a safety net setting. Particularly for patients lacking adequate health literacy, English proficiency, and/or the technological know-how or interest to access and interpret their personal health information, there is potential for caregivers to use patient portals to improve their ability to interpret health information, coordinate care, and assist with medical decision making [22]. Past studies have found high interest among patients in sharing their personal health information with caregivers [42,43] and among caregivers in accessing patient health information through health technology such as portals [44,45]. Caregivers in our study expanded on the utility of having access to patient health information to describe a deeper role in caregiving, noting their role as interpreters of this information, guides in decision making and behavior change, and advocates in ensuring quality of care. Our findings illustrate the need for safety net health systems implementing patient portals to address caregiver needs—including strategies for formal proxy processes for patients to officially grant others access to their portal account as well as awareness of informal sharing of username/passwords between family members that is also likely to occur. Because the caregivers in our sample were already highly involved in health care management tasks, further research is needed to understand perspectives on the levels of caregiver access to personal health information [40,42].

Limitations

Because our small study examined the in-depth perspectives of patients receiving care from one urban safety net hospital, our findings are likely not generalizable to patients receiving care from larger health systems or networks. Furthermore, because our eligibility criteria required participants to speak English and express at least some interest in using the Internet to manage their health, our findings may not be generalizable to those with limited English proficiency or needing even more basic computer or technology training. Finally, our study did not incorporate perspectives on the actual usability of a live portal website. Instead, we focused on gaining in-depth information about the barriers and facilitators to portal use in advance of its rollout to support a patient-centered approach to implementing our portal system-wide within the San Francisco Health Network.

Conclusions

Our findings indicate that interest in using patient portals may not always match the technological proficiency of more vulnerable patients. This indicates a need for safety net health systems or other social service providers (eg, library, adult literacy classes) to provide training not only in portal use, but also in equipping patients with the basic computer and health literacy to effectively use a portal. To address patient concerns, it is important for health care systems implementing portals to assess the potential effects of the replacement of in-person or verbal communication resulting from portal use and establish high levels of online security.

From a national perspective, our findings suggest that widespread EHR and portal implementation may be hindered by patient engagement challenges in the coming years, especially with respect to health literacy and language proficiency status. Although incentives to promote meaningful use have been successful at driving health systems to implement health information technology, these standards do not guarantee that safety net health systems will adopt the newest or most accessible technologies on the market [46,47] or that patient portals will be accessible or useful to all patients, especially those who may face additional limitations in literacy and technology experience. Addressing health literacy and other barriers may best be achieved through patient-centered approaches to the adoption of health information technology at the planning, implementation, and evaluation stages [48]. If implemented with patient perspectives in mind, patient portals have the potential to be a convenient and effective way to improve self-management and quality of care for patients and caregivers receiving care from safety net settings.

Acknowledgments

Funding for this study comes from the Agency for Healthcare Research & Quality (K99 HS022408, R24HS022047, 1K08HS022561), NIH/NIDDK (DK092924-01), NIH/NIMHD (1P60MD006902), NIH/NLM (G08LM012166), and NIH/NCATS (KL2TR000143).

Authors' Contributions

LT contributed to protocol design, data acquisition, data analysis, and drafting of the manuscript. US, DS, JDR, NR, and RP guided the analyses and edited and approved the manuscript. CRL conceptualized the study and contributed to protocol design, data acquisition, data analysis, and editing and approval of the manuscript. All authors have fulfilled the criteria for authorship established by the International Committee of Medical Journal Editors and approved submission of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview Guide.

[PDF File (Adobe PDF File), 318KB - jmir_v17i12e275_app1.pdf]

References

1. Vol Title XIII of Division A and Title IV of Division B of the American Recovery and Reinvestment Act. Washington, DC: Office of the National Coordinator for Health Information; 2009 Feb 18. Health Information Technology for Economic and Clinical Health (HITECH) Act URL: https://www.healthit.gov/sites/default/files/hitech_act_excerpt_from_arra_with_index.pdf [accessed 2015-11-16] [WebCite Cache ID 6d7XdrRUw]
2. HealthIT.gov. EHR incentives & certification: meaningful use definitions & objectives URL: http://www.healthit.gov/providers-professionals/meaningful-use-definition-objectives [accessed 2015-06-20] [WebCite Cache ID 6ZREeafHS]
3. Centers for Medicare and Medicaid Services. Electronic Health Records Incentive Program-Stage 2. Federal Register 2012 Sep 04;77(171):53968-54162.
4. Urowitz S, Wiljer D, Dupak K, Kuehner Z, Leonard K, Lovrics E, et al. Improving diabetes management with a patient portal: a qualitative study of diabetes self-management portal. J Med Internet Res 2012;14(6):e158 [FREE Full text] [doi: 10.2196/jmir.2265] [Medline: 23195925]
5. Dhanireddy S, Walker J, Reisch L, Oster N, Delbanco T, Elmore JG. The urban underserved: attitudes towards gaining full access to electronic medical records. Health Expect 2014 Oct;17(5):724-732. [doi: 10.1111/j.1369-7625.2012.00799.x] [Medline: 22738155]
6. Winkelman WJ, Leonard KJ, Rossos PG. Patient-perceived usefulness of online electronic medical records: employing grounded theory in the development of information and communication technologies for use by patients living with chronic illness. J Am Med Inform Assoc 2005;12(3):306-314 [FREE Full text] [doi: 10.1197/jamia.M1712] [Medline: 15684128]
7. Osborn CY, Mayberry LS, Wallston KA, Johnson KB, Elasy TA. Understanding patient portal use: implications for medication management. J Med Internet Res 2013;15(7):e133 [FREE Full text] [doi: 10.2196/jmir.2589] [Medline: 23823974]
8. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]
9. Osborn CY, Mayberry LS, Mulvaney SA, Hess R. Patient web portals to improve diabetes outcomes: a systematic review. Curr Diab Rep 2010 Dec;10(6):422-435 [FREE Full text] [doi: 10.1007/s11892-010-0151-1] [Medline: 20890688]
10. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011 May 1;18(3):318-321 [FREE Full text] [doi: 10.1136/jamia.2010.006015] [Medline: 21262921]
11. Lau M, Campbell H, Tang T, Thompson DJ, Elliott T. Impact of patient use of an online patient portal on diabetes outcomes. Can J Diabetes 2014 Feb;38(1):17-21. [doi: 10.1016/j.jcjd.2013.10.005] [Medline: 24485208]
12. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]
13. Lyles CR, Harris LT, Jordan L, Grothaus L, Wehnes L, Reid RJ, et al. Patient race/ethnicity and shared medical record use among diabetes patients. Med Care 2012 May;50(5):434-440. [doi: 10.1097/MLR.0b013e318249d81b] [Medline: 22354209]
14. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, Lopez A, et al. The literacy divide: health literacy and the use of an internet-based patient portal in an integrated health system-results from the diabetes study of northern California (DISTANCE). J Health Commun 2010;15 Suppl 2:183-196 [FREE Full text] [doi: 10.1080/10810730.2010.499988] [Medline: 20845203]
15. Roblin DW, Houston TK, Allison JJ, Joski PJ, Becker ER. Disparities in use of a personal health record in a managed care organization. J Am Med Inform Assoc 2009;16(5):683-689 [FREE Full text] [doi: 10.1197/jamia.M3169] [Medline: 19567790]
16. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
17. Ancker JS, Barrón Y, Rockoff ML, Hauser D, Pichardo M, Szerencsy A, et al. Use of an electronic patient portal among disadvantaged populations. J Gen Intern Med 2011 Oct;26(10):1117-1123 [FREE Full text] [doi: 10.1007/s11606-011-1749-y] [Medline: 21647748]

18. Goel MS, Brown TL, Williams A, Hasnain-Wynia R, Thompson JA, Baker DW. Disparities in enrollment and use of an electronic patient portal. J Gen Intern Med 2011 Oct;26(10):1112-1116 [FREE Full text] [doi: 10.1007/s11606-011-1728-3] [Medline: 21538166]
19. Zikmund-Fisher BJ, Exe NL, Witteman HO. Numeracy and literacy independently predict patients' ability to identify out-of-range test results. J Med Internet Res 2014;16(8):e187 [FREE Full text] [doi: 10.2196/jmir.3241] [Medline: 25135688]
20. Zickmund SL, Hess R, Bryce CL, McTigue K, Olshansky E, Fitzgerald K, et al. Interest in the use of computerized patient portals: role of the provider-patient relationship. J Gen Intern Med 2008 Jan;23 Suppl 1:20-26 [FREE Full text] [doi: 10.1007/s11606-007-0273-6] [Medline: 18095039]
21. Zarcadoolas C, Vaughon WL, Czaja SJ, Levy J, Rockoff ML. Consumers' perceptions of patient-accessible electronic medical records. J Med Internet Res 2013;15(8):e168 [FREE Full text] [doi: 10.2196/jmir.2507] [Medline: 23978618]
22. National Alliance for Caregiving, AARP. Caregiving in the US 2009. 2009 Nov. URL: http://www.caregiving.org/data/Caregiving_in_the_US_2009_full_report.pdf [accessed 2015-06-20] [WebCite Cache ID 6ZREnnbmt]
23. The Institute for Family Health. 2009 Oct 13. The Institute for Family Health and the National Library of Medicine help New York patients find health information online URL: http://www.institute.org/news/the-institute-for-family-health-and-the-national-library-of-medicinethe-institute-for-family-health-and-the-national-library-of-medicine [accessed 2015-06-19] [WebCite Cache ID 6ZREqIssA]
24. Kanaan S. Safety Net Providers Bring Patients Online: Lessons from Early Adopters. Oakland, CA: California Healthcare Foundation; 2009. URL: http://www.chcf.org/~/media/MEDIA%20LIBRARY%20Files/PDF/S/PDF%20SafetyNetPatientsOnline.pdf [accessed 2015-06-20] [WebCite Cache ID 6ZREuvY4E]
25. López A, Patterson AM, John V, Sarkar U. Electronic health record implementation in outpatient safety-net settings in California. J Health Care Poor Underserved 2012 Nov;23(4):1421-1430. [doi: 10.1353/hpu.2012.0154] [Medline: 23698659]
26. Schickedanz A, Huang D, Lopez A, Cheung E, Lyles CR, Bodenheimer T, et al. Access, interest, and attitudes toward electronic communication for health care among patients in the medical safety net. J Gen Intern Med 2013 Jul;28(7):914-920 [FREE Full text] [doi: 10.1007/s11606-012-2329-5] [Medline: 23423453]
27. McInnes DK, Solomon JL, Shimada SL, Petrakis BA, Bokhour BG, Asch SM, et al. Development and evaluation of an internet and personal health record training program for low-income patients with HIV or hepatitis C. Med Care 2013 Mar;51(3 Suppl 1):S62-S66. [doi: 10.1097/MLR.0b013e31827808bf] [Medline: 23407015]
28. Chew LD, Bradley KA, Boyko EJ. Brief questions to identify patients with inadequate health literacy. Fam Med 2004 Sep;36(8):588-594 [FREE Full text] [Medline: 15343421]
29. Venkatesh V, Morris M, Davis G, David F. User acceptance of information technology: toward a unified view. MIS Quarterly 2003;27(3):425-478.
30. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Quarterly 2012;36(1):157-178.
31. Kahlke R. Generic Qualitative Approaches: Pitfalls and Benefits of Methodological Mixology. International Journal of Qualitative Methods 2014;13:37-49.
32. Thorne S, Kirkham SR, MacDonald-Emes J. Interpretive description: a noncategorical qualitative alternative for developing nursing knowledge. Res Nurs Health 1997 Apr;20(2):169-177. [Medline: 9100747]
33. Atlas.ti Qualitative Data Analysis. Berlin: Scientific Software Development; 2013. URL: http://atlasti.com/wp-content/uploads/2014/05/atlasti_v7_manual_201312.pdf?q=/uploads/media/atlasti_v7_manual_201312.pdf [accessed 2015-11-17] [WebCite Cache ID 6d7amaEfN]
34. Miles M, Huberman M. Qualitative Data Analyses: An Expanded Sourcebook. Thousand Oaks, CA: Sage Publications; 1994.
35. Pope C, Mays N. Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. BMJ 1995 Jul 1;311(6996):42-45 [FREE Full text] [Medline: 7613329]
36. Hsu J, Huang J, Kinsman J, Fireman B, Miller R, Selby J, et al. Use of e-Health services between 1999 and 2002: a growing digital divide. J Am Med Inform Assoc 2005;12(2):164-171 [FREE Full text] [doi: 10.1197/jamia.M1672] [Medline: 15561786]
37. Hassol A, Walker JM, Kidder D, Rokita K, Young D, Pierdon S, et al. Patient experiences and attitudes about access to a patient electronic health care record and linked web messaging. J Am Med Inform Assoc 2004;11(6):505-513 [FREE Full text] [doi: 10.1197/jamia.M1593] [Medline: 15299001]
38. Wade-Vuturo AE, Mayberry LS, Osborn CY. Secure messaging and diabetes management: experiences and perspectives of patient portal users. J Am Med Inform Assoc 2013 May 1;20(3):519-525 [FREE Full text] [doi: 10.1136/amiajnl-2012-001253] [Medline: 23242764]
39. Smith SG, O'Conor R, Aitken W, Curtis LM, Wolf MS, Goel MS. Disparities in registration and use of an online patient portal among older adults: findings from the LitCog cohort. J Am Med Inform Assoc 2015 Apr 25:888-895. [doi: 10.1093/jamia/ocv025] [Medline: 25914099]
40. Goel MS, Brown TL, Williams A, Cooper AJ, Hasnain-Wynia R, Baker DW. Patient reported barriers to enrolling in a patient portal. J Am Med Inform Assoc 2011 Dec;18 Suppl 1:i8-12 [FREE Full text] [doi: 10.1136/amiajnl-2011-000473] [Medline: 22071530]

41. Kim E, Stolyar A, Lober WB, Herbaugh AL, Shinstrom SE, Zierler BK, et al. Challenges to using an electronic personal health record by a low-income elderly population. J Med Internet Res 2009;11(4):e44 [FREE Full text] [doi: 10.2196/jmir.1256] [Medline: 19861298]
42. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. Ann Intern Med 2011 Dec 20;155(12):811-819 [FREE Full text] [doi: 10.7326/0003-4819-155-12-201112200-00003] [Medline: 22184688]
43. Zulman DM, Nazi KM, Turvey CL, Wagner TH, Woods SS, An LC. Patient interest in sharing personal health record information: a web-based survey. Ann Intern Med 2011 Dec 20;155(12):805-810. [doi: 10.7326/0003-4819-155-12-201112200-00002] [Medline: 22184687]
44. Zulman DM, Piette JD, Jenchura EC, Asch SM, Rosland A. Facilitating out-of-home caregiving through health information technology: survey of informal caregivers' current practices, interests, and perceived barriers. J Med Internet Res 2013;15(7):e123 [FREE Full text] [doi: 10.2196/jmir.2472] [Medline: 23841987]
45. Mishuris RG, Stewart M, Fix GM, Marcello T, McInnes DK, Hogan TP, et al. Barriers to patient portal access among veterans receiving home-based primary care: a qualitative study. Health Expect 2014 May 12:e1. [doi: 10.1111/hex.12199] [Medline: 24816246]
46. Hart JT. The inverse care law. Lancet 1971 Feb 27;1(7696):405-412. [Medline: 4100731]
47. Schillinger D. Literacy and health communication: reversing the 'inverse care law'. Am J Bioeth 2007 Nov;7(11):15-8; discussion W1. [doi: 10.1080/15265160701638553] [Medline: 18027291]
48. Brach C, Keller D, Hernandez LM, Baur C, Parker R, Dreyer B, et al. Ten Attributes of Health Literate Health Care Organizations. Washington, DC: Institute of Medicine; 2012. URL: http://iom.edu/~/media/Files/Perspectives-Files/2012/Discussion-Papers/BPH_Ten_HLit_Attributes.pdf [accessed 2015-06-20] [WebCite Cache ID 6ZRFI5rvj]

Abbreviations

COPD: chronic obstructive pulmonary disease
EHR: electronic health record
GMC: General Medicine Clinic
HITECH: Health Information Technology for Economic and Clinical Health
SFGH: San Francisco General Hospital

Edited by N Zary; submitted 20.06.15; peer-reviewed by L Mayberry, C Kruse; comments to author 17.07.15; revised version received 14.08.15; accepted 09.10.15; published 03.12.15.

Please cite as:
Tieu L, Sarkar U, Schillinger D, Ralston JD, Ratanawongsa N, Pasick R, Lyles CR
Barriers and Facilitators to Online Portal Use Among Patients and Caregivers in a Safety Net Health Care System: A Qualitative Study
J Med Internet Res 2015;17(12):e275
URL: http://www.jmir.org/2015/12/e275/
doi:10.2196/jmir.4847
PMID:26681155

Original Paper

“Why Didn’t it Work?” Lessons From a Randomized Controlled Trial of a Web-based Personally Controlled Health Management System for Adults with Asthma

Annie YS Lau[1], PhD; Amaël Arguel[2], PhD; Sarah Dennis[3], PhD; Siaw-Teng Liaw[4], MBBS, PhD; Enrico Coiera[1], MBBS, PhD

[1]Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Sydney, Australia

[2]ARC-SRI Science of Learning Research Centre, Macquarie University, Sydney, Australia

[3]Discipline of Physiotherapy, University of Sydney, Sydney, Australia

[4]Centre for Primary Health Care and Equity, School of Public Health & Community Medicine, UNSW Medicine, University of New South Wales, Sydney, Australia

Corresponding Author:
Annie YS Lau, PhD
Centre for Health Informatics
Australian Institute of Health Innovation
Macquarie University
Level 6 Talavera Rd Macquarie University
Sydney, 2109
Australia
Phone: 61 431599890
Fax: 61 8088 6234
Email: annie.lau@mq.edu.au

Abstract

Background: Personally controlled health management systems (PCHMS), which may include a personal health record (PHR), health management tools, and information resources, have been advocated as a next-generation technology to improve health behaviors and outcomes. There have been successful trials of PCHMS in various health settings. However, there is mixed evidence for whether consumers will use these systems over the long term and whether they ultimately lead to improved health outcomes and behaviors.

Objective: The aim was to test whether use of a PCHMS by consumers can increase the uptake or updating of a written asthma action plan (AAP) among adults with asthma.

Methods: A 12-month parallel 2-group randomized controlled trial was conducted. Participants living with asthma were recruited nationally in Australia between April and August 2013, and randomized 1:1 to either the PCHMS group or control group (online static educational content). The primary outcome measure was possession of an up-to-date written AAP poststudy. Secondary measures included (1) utilizing the AAP; (2) planned or unplanned visits to a health care professional for asthma-related concerns; (3) severe asthma exacerbation, inadequately controlled asthma, or worsening of asthma that required a change in treatment; and (4) number of days lost from work or study due to asthma. Ancillary analyses examined reasons for adoption or nonadoption of the intervention. Outcome measures were collected by online questionnaire prestudy, monthly, and poststudy.

Results: A total of 330 eligible participants were randomized into 1 of 2 arms (intervention: n=154; control: n=176). Access to the PCHMS was not associated with a significant difference in any of the primary or secondary outcomes. Most participants (80.5%, 124/154) did not access the intervention or accessed it only once.

Conclusions: Despite the intervention being effective in other preventive care settings, system use was negligible and outcome changes were not seen as a result. Consumers must perceive the need for assistance with a task and assign priority to the task supported by the eHealth intervention. Additionally, the cost of adopting the intervention (eg, additional effort, time spent learning the new system) must be lower than the benefit. Otherwise, there is high risk consumers will not adopt the eHealth intervention.

Trial Registration: Australian New Zealand Clinical Trials Registry (ANZCTR): ACTRN12612000716864; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=362714 (Archived by WebCite® at http://www.webcitation.org/6dMV6hg4A)

(J Med Internet Res 2015;17(12):e283) doi:10.2196/jmir.4734

KEYWORDS

asthma; self-management; personal health record; personally controlled health management system; eHealth; Internet; intervention; adult

Introduction

Personally controlled health management systems (PCHMS), which may include a personal health record (PHR), health management tools, and information resources, have been advocated as a next-generation technology to improve health behaviors and outcomes [1]. Trials of PCHMS have been undertaken in various health settings, including in vitro fertilization [2], hypertension [3], diabetes [4,5], influenza vaccination [6,7], sexually transmitted infection [8], medication accuracy [9], breast cancer management [10], and physical and emotional well-being [11,12]. However, there is mixed evidence for whether consumers will use these systems over the long term and whether they ultimately lead to improved health outcomes and behaviors [2-7,9-12].

Lack of engagement in digital interventions is a common phenomenon [13,14] and the reasons for adoption or nonadoption remain underreported. An early analysis of the reasons that led to the abandonment of a national PHR in the United Kingdom concluded that unless a system aligned closely with people's attitudes, practices, information needs, and preexisting health services, then the risk it will not be adopted is substantial [15]. More recent analyses suggest that for chronic illnesses, PCHMS work best when there is a feedback loop between monitoring in the PHR and behaviors that could be self-managed by a consumer [16].

In this study, we examine how effective a PCHMS is in encouraging adults with a chronic condition—asthma—to obtain a written asthma action plan (AAP) from their primary care practitioner over a 12-month period. We also explored the reasons that underlay the adoption or nonadoption of the intervention.

Asthma and the Written Asthma Action Plan

Asthma is a chronic condition [17,18]. The prevalence of asthma is significant; worldwide, the number of people suffering from asthma is approximately 300 million [17]. According to a Cochrane review, one of the most efficient tools that patients can use to manage their asthma is a written AAP [19]. The AAP is a set of instructions prepared with a health care professional that helps recognize signs that asthma is worsening, indicates which medication to use, or provides nonmedication strategies to keep asthma under control [20]. Written AAPs are individualized documents that must be updated (eg, once a year) by a clinician to match the evolution of the individual's asthma condition [20].

When properly used, written AAPs are associated with fewer visits to the emergency department with an asthma exacerbation, fewer hospital admissions, better lung function, and an overall improvement of asthma symptoms [19]. Although having an up-to-date AAP is highly recommended, written AAPs are widely underused among adults; only 1 in 5 patients actually possess an up-to-date and usable AAP [21,22]. Most initiatives to improve the uptake of the written AAP have been targeted at health care professionals. Few have targeted patients.

This randomized controlled trial (RCT) is designed to test whether a PCHMS, tailored to help adults with asthma, would increase their rate of obtaining or updating a written AAP from a health care professional and whether this would lead to an improvement in asthma control.

Hypotheses

Compared to participants allocated to the control group (ie, static online educational page), we hypothesized that those using a PCHMS are

1. More likely to obtain or update a written AAP;
2. More likely to make planned visits to a health care professional for asthma;
3. Less likely to make unplanned visits to a health care professional for asthma; and
4. Less likely to experience (1) severe asthma exacerbation, (2) inadequately controlled asthma, (3) worsening of asthma that requires a change in treatment, or (4) days lost from work or study due to asthma.

Methods

Details on participants, recruitment strategy, intervention description, data collection, ethical considerations, and study procedure are described in the study protocol [23]. Utilization of AAP was defined by participant self-report to the questions "During the study, when you experienced an asthma exacerbation, did you use your written asthma action plan?" and "How often did you use your AAP during the study?"

Trial Design

In this parallel 2-arm RCT, participants were stratified by gender and level of asthma severity (intermittent vs persistent), and randomized 1:1 to have immediate access to the PCHMS or to control.

Participant recruitment took place between April and August 2013. All individuals who expressed an interest were assessed with a 5-minute online eligibility questionnaire. Eligible individuals were then invited to complete a 10- to 15-minute prestudy questionnaire. Participants in both arms continued to receive usual care from their health services and were surveyed monthly for asthma symptoms, asthma exacerbation, asthma control, and other competing priorities, and followed up with a 10- to 15-minute poststudy questionnaire between April and June 2014.

Participants and Setting

Eligible participants were adults (aged 18 years and older) living in Australia diagnosed with asthma, who had at least monthly access to the Internet and email, and had sufficient English language skills.

Control Group

On completion of the prestudy questionnaire, participants who had been randomly allocated to the control arm were redirected to a static webpage with links to patient websites (eg, the Asthma Foundation, HealthInsite, myDR) that provided educational information on asthma. They were advised that they would be contacted to complete monthly surveys to elicit their asthma status during the study and would receive a follow-up questionnaire on conclusion of the study.

Intervention Group

Full details of the Healthy.me Web-based PCHMS are described elsewhere [11,12,23]. During the study, Healthy.me provided participants with evidence-based information about asthma, the importance of a written AAP, and ways of obtaining a plan from a health care professional. Additionally, participants received monthly email reminders about the various interactive features of Healthy.me (eg, forum, poll, PHR).

An expert steering group was formed to tailor educational content for patients with asthma and to customize the interactive features of Healthy.me to deliver this content over the 12 months. Three asthma "journeys" were developed, providing evidence-based material to consumers about the written AAP. A usability study with 10 individuals was conducted and all major usability issues associated with the content and the intervention were addressed before commencing the study.

Theoretical Framework of the Intervention

A review of online interventions found that those built on a theoretical framework demonstrated greater efficacy [24]. The Health Belief Model (HBM) [25], a prominent model of behavioral change, was used to guide the design of the 3 asthma journeys. More details are available in Multimedia Appendix 1.

There are strong theoretical reasons why the PCHMS features drive behavioral change:

1. The online appointment booking service, embedded within health service information descriptions, allows consumers to turn information into action in keeping with the "cue to action" elements of the HBM [25].
2. Social features (eg, polls and forums), which allow individuals to connect with others and observe social norms on health behaviors, are designed according to principles of social cognitive theory [26].
3. PHRs, which facilitate self-management and self-awareness, are related to the principle of increasing self-efficacy.
4. The journey model, which allows stages of change described in a step-by-step manner, is congruent with the Theory of Transtheoretical Change [27].

Analysis Method

All primary, secondary, and ancillary analyses are outlined in the published protocol [23]. Sample size calculation and expected effect size are also documented in the protocol [23]. No major changes from the protocol were introduced during study execution. Statistical significance was defined as a *P* value of less than .05 (2-sided test). Effect sizes are reported with 95% confidence intervals. Data were analyzed using SPSS version 20.

Primary Analysis

The intention-to-treat principle was followed in the primary analysis. Missing values were managed by the last observation carried forward (LOCF) imputation procedure [28]. The Pearson chi-square test was used to identify any significant difference between the proportion of participants in the control and intervention groups who reported having obtained (or updated) a written AAP during the study.

Binary logistic regression was employed to adjust for potential confounding factors or differences in baseline characteristics that were expected to be predictive of the outcome, including age, gender, past possession of a written AAP, smoking status, medications used for asthma, and past visits to a health care professional for asthma concerns [29].

Secondary Analyses

A complete case analysis of secondary outcomes was also conducted using the data of those who completed the poststudy questionnaire and the Pearson chi-square test to identify any significant difference between intervention and control groups.

A comparison was made of the proportion of patients in the intervention and control groups who reported experiencing at least one of the following episodes during the study:

1. Severe asthma exacerbation (as indicated in the Official American Thoracic Society/European Respiratory Society Statement on Asthma Control and Exacerbations) [30];
2. Inadequate asthma control (as measured by Asthma Control Questionnaire [ACQ] score of ≥1.5 in that month) [31];
3. Worsening of asthma that required treatment changes (as measured by a decrease in ACQ score of ≥0.5 between 2 consecutive months) [32]; and
4. Missing one or more days from work or study due to asthma.

Ancillary Analyses

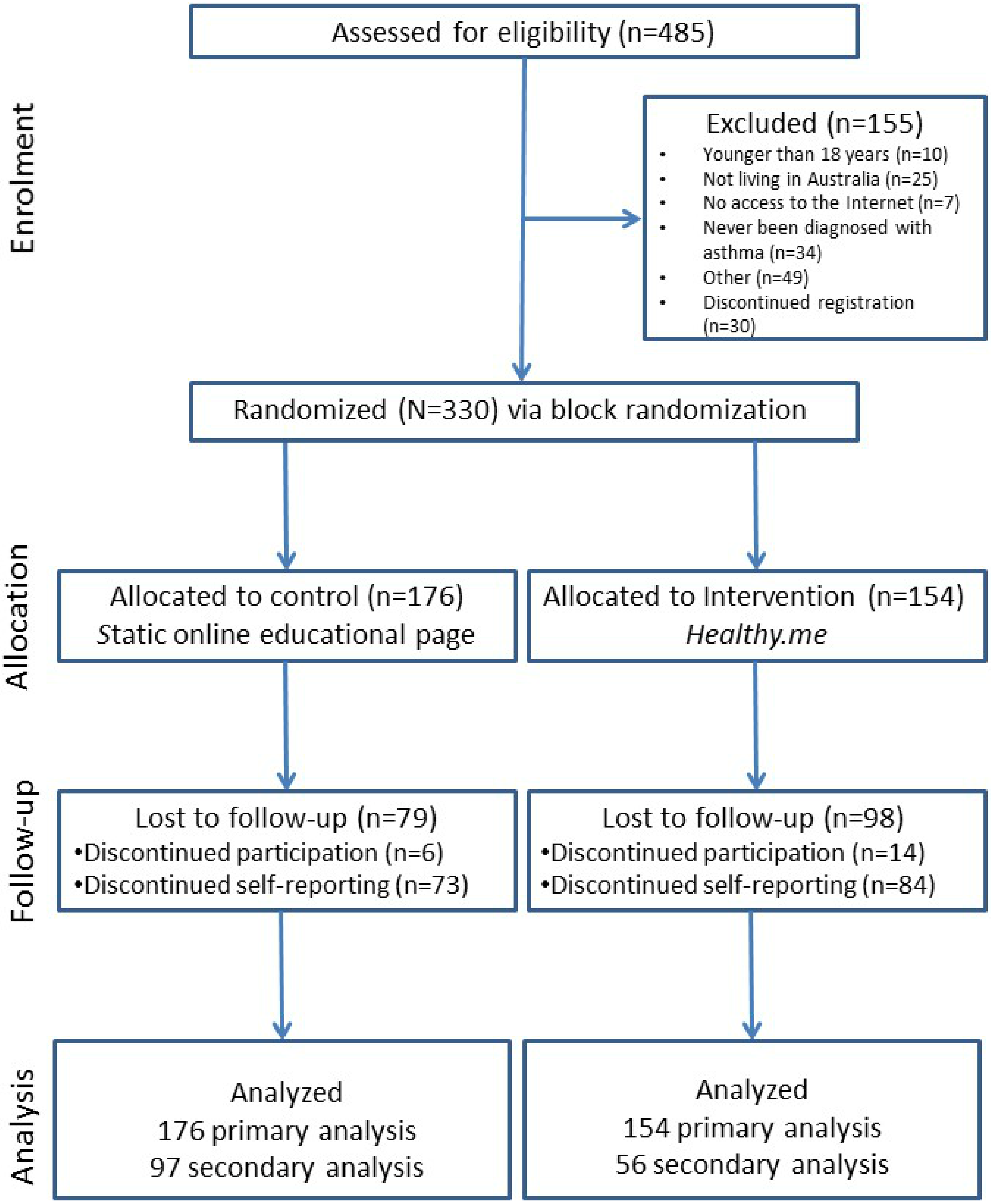
Ancillary analyses were conducted to examine reasons for adoption or nonadoption of the intervention. These were conducted using the data of those who completed the poststudy questionnaire or at least one monthly questionnaire. Participant engagement with the intervention was measured via system logs and their perception of the intervention was measured by the Technology Acceptance Model (TAM) instrument [33]. Outcome measures included reasons for obtaining (or not obtaining) a written AAP, participant competing priorities, responses to TAM, and usage of the PCHMS. All measures were reported using descriptive statistics and illustrated with written feedback collected in the poststudy questionnaires. Any recurring patterns or themes reported were emergent from the written feedback from participants. Participant quotes were reported with no alterations.

Results

Participant Recruitment, Flow, and Exclusions

Recruitment was conducted over a period of 5 months between April and August 2013, during which 485 participants were assessed for eligibility (Figure 1). Recruitment was complete in August 2013 and follow-up conducted between April and June 2014. In all, 330 participants were assessed eligible and randomized (intervention: n=154; control: n=176). No participants with available data were excluded from the analyses. No harm or unintended effects were reported by participants during the study.

Figure 1. Participant flowchart.



Baseline Data

Baseline characteristics were similar for all allocated participants, participants lost to follow-up, and remaining participants (Table 1). The majority of participants were female (control: 79.5%, 140/176; intervention: 80.5%, 124/154), in their late thirties / early forties (control: mean 39, SD 13 years; intervention: mean 40, SD 14 years), and were very familiar with the Internet and social networking sites such as Facebook and Twitter. Asthma-related characteristics, such as smoking status, use of asthma medication, and contact with health care professional for asthma in the past 12 months, were similar to rates identified in a national survey of adults with asthma [22]. However, their rate of possessing a written AAP was higher than the national rate (this study: 38.8%, 128/330 vs Australian Centre for Asthma Monitoring: 21.3% [22]). For those with a written AAP at prestudy, 75.8% (97/128) could not recall when they last obtained or updated it.

Table 1. Baseline characteristics of all participants and those lost to follow-up.

Baseline characteristic	All participants		Participants lost to follow-up		Remaining participants	
	Control (n=176)	Intervention (n=154)	Control (n=79)	Intervention (n=98)	Control (n=97)	Intervention (n=56)
Female, n (%)	140 (79.5)	124 (80.5)	59 (75)	80 (82)	81 (84)	44 (79)
Age (years), mean (SD)	39 (13)	40 (14)	37 (12)	36 (12)	41 (14)	46 (14)
Has written AAP (before study), n (%)	71 (40.3)	57 (37.0)	28 (35)	37 (38)	43 (44)	20 (36)
Visited health care professional for asthma in past 12 months, n (%)	142 (80.7)	133 (86.4)	71 (90)	83 (85)	71 (73)	50 (89)
Smoking status, n (%)	16 (9.1)	13 (8.4)	10 (17)	12 (12)	6 (6)	1 (2)
Preventer use in the past 12 months, n (%)[a]	87 (49.4)	78 (50.6)	34 (43)	49 (50)	53 (54)	29 (52)
Reliever use in the past 12 months, n (%)[b]	169 (96.0)	149 (96.8)	76 (96)	93 (95)	93 (96)	56 (100)
Symptom controller use in the past 12 months, n (%)[c]	7 (4.0)	7 (4.5)	4 (5)	5 (5)	3 (3)	2 (4)
Visit social networking sites (eg, Facebook, Twitter) several times a day, n (%)	131 (74.4)	99 (64.3)	67 (85)	66 (67)	64 (66)	33 (59)
Never used the Internet to find health information, n (%)	4 (2.3)	8 (5.2)	2 (3)	6 (6)	2 (2)	2 (4)

[a] Preventer use: Flixotide, Pulmicort, Qvar, Alvesco, Leukotriene, Singulair, Cromones, Intal, Tilade, Xolair (Omalizumab).

[b] Reliever use: Ventolin, Asmol, Epaq, Airomir, Bricanyl, Atrovent.

[c] Symptom controller use: Serevent, Oxis, Fovadile.

Numbers Analyzed

Analyses of the primary outcome (possession of a written AAP at poststudy) was conducted by intention-to-treat using the data of all 330 allocated participants and the 153 participants who completed the poststudy questionnaire.

We did not apply the intention-to-treat principle to secondary and ancillary outcomes due to the availability of data for analyses. Analyses of secondary outcomes relating to use of the AAP and visits to a health care professional were conducted using the data of 153 participants who completed the poststudy questionnaire. Other study outcomes (ie, asthma exacerbation, asthma control, worsening of asthma, and loss days from work or study) were conducted using the data of 242 participants who completed at least one monthly questionnaire.

Analysis of Primary Outcome

Analysis of the primary outcome is outlined in Table 2 (more details in Multimedia Appendix 2). There were no significant differences in the proportion of participants who reported having a written AAP poststudy between the intervention and control groups (all participants: $\chi^2_1=0.6$, $P=.43$; all participants with imputation method: $\chi^2_1=0.4$, $P=.52$; remaining participants at study end: $\chi^2_1=0.9$, $P=.36$).

Table 2. Analysis of primary outcome by study group (for all participants and remaining participants).

Analysis	n	Has written AAP (poststudy)			Has written AAP (poststudy LOCF[a])		
		n (%)	χ^2_1	*P*	n (%)	χ^2_1	*P*
All participants			0.6	.43		0.4	.52
Control	176	38 (22)			66 (38)		
Intervention	154	27 (18)			64 (42)		
Remaining participants							
Control	97	38 (39)	0.9	.36			
Intervention	56	27 (48)					

[a] LOCF: last observation carried forward (imputation method to address missing data).

Binary logistic regression was adjusted for differences in baseline characteristics and potential confounding factors that might influence the primary outcome measure. Only one independent variable made a statistically significant contribution to the regression model: possession of written AAP at prestudy (χ^2_{22}=299.6, *P*<.001). Allocation to the intervention group (OR 0.43, 95% CI 0.15-1.23) did not contribute a significant effect to the proportion of participants possessing a written AAP at poststudy.

Analysis of Secondary Outcomes

Analyses of secondary outcomes are presented in Table 3 (more details in Multimedia Appendix 2). There were no statistically significant differences in the proportion of participants between the control and intervention groups for any secondary outcome.

Table 3. Analyses of secondary outcomes by study group (for remaining participants).

Analysis	Participants, n (%)		χ^2_1	*P*
	Control	Intervention		
Completed poststudy questionnaire (control: n=97; intervention: n=56)				
Used AAP more than once during study	20 (21)	11 (20)	0.02	.89
Visited health care professional for nonemergency asthma	58 (60)	36 (64)	0.1	.71
Visited health care professional for emergency/urgent asthma	42 (43)	24 (43)	0.003	.96
Visited emergency department for emergency/unplanned asthma	15 (15)	10 (18)	0.03	.88
Visited GP or respiratory physician for emergency/unplanned asthma	35 (36)	20 (36)	0.002	.96
Completed ≥1 monthly questionnaire (control: n=145; intervention: n=97)				
Severe asthma exacerbation at least once during study	62 (43)	35 (36)	0.8	.37
Asthma inadequately controlled at least once during study (as measured by ACQ score ≥1.5)	137 (94)	87 (90)	1.3	.25
Worsening of asthma that requires a change in treatment (as measured by a decrease of ≥0.5 in ACQ score between 2 consecutive months)	77 (53)	44 (45)	1.1	.29
Lost days from work or school due to asthma during study	61 (42)	33 (34)	1.3	.26

Ancillary Analyses

Reasons for Not Obtaining a Written Asthma Action Plan

Participant reasons for not obtaining or updating a written AAP during the study are outlined in Table 4. Among control participants, the most frequently cited reason was their “lack of awareness of the plan” (31%, 18/59). Whereas, the reason for PCHMS participants was “none of the above” (33%, 9/27), which included other knowledge, motivation, or belief-related reasons that were not anticipated (such as perceiving the plan to be “irrelevant,” lack of importance placed on asthma, or other life and health priorities which competed for their attention).

Table 4. Reasons for not obtaining/updating a written AAP by study group.[a]

Reasons	Participants, n (%)	
	Control (n=59)	Intervention (n=27)
Knowledge, motivation, or belief-related		
I did not know about the existence of AAPs	18 (31)	8 (30)
I do not believe that a written AAP could be useful to me	10 (17)	4 (15)
I lacked the motivation to get a written AAP	9 (15)	2 (7)
I do not know where to get it	7 (12)	1 (4)
I think the written AAP could be difficult to use	2 (3)	3 (11)
Other reasons		
I did not visit a doctor during the study	13 (22)	6 (22)
I lacked the time to get a written AAP	10 (17)	2 (7)
I simply forgot	9 (15)	2 (7)
It was inconvenient to get it	1 (2)	0
None of the above, please specify[b]	12 (20)	9 (33)

[a] Participants could select more than one reason.

[b] Reasons such as perceiving the plan to be "irrelevant," lack of importance placed on asthma, or other life and health priorities which competed for their attention.

For those who provided further explanation in the poststudy questionnaires, a variety of reasons for not obtaining (or updating) a written AAP were offered.

Participants feeling comfortable with a verbal plan or their own experience in self-management:

> *My doctor and I have discussed this in detail. No written plan required.* [Participant ID 13; accessed PCHMS: once]
>
> *I am familiar with the steps in a written plan and follow these principles; however, I don't require an actual hard copy of one, as I am confident in my self-management. Also, the last doctor who tried to force one upon me did not even try to understand my asthma or lifestyle, rather insulting me rude [rudely] and thinking that a generic plan (that included medicine that I do not respond to) was the only way to go. I'm sure he was an exception, but I'm honestly fine with the way I manage my asthma, and when I ask GPs about my medication, it's rare that there's anything new going on.* [Participant ID 70; accessed PCHMS: once]

The poor experience they had with previous written plans / past health care professionals:

> *I got one a while ago and it was a tick and flick from a drug company and I felt it was useless—gave me nothing more than I know now.* [Participant ID 38; accessed PCHMS: once]
>
> *I have never received one for me though I am a severe asthmatic. My child has received one for whenever he is sick and ends up in hospital. But we have no action plan for either of us for what to do on a normal day and we are feeling unwell with signs and symptoms of asthma.* [Participant ID 80; accessed PCHMS: once]

Discouragement by health care professionals:

> *It's never been offered by doctor.* [Participant ID 74; accessed PCHMS: zero times]
>
> *Doctor told me not to bother.* [Participant ID 3; accessed PCHMS: zero times]

Competing priorities experienced during the study:

> *I really didn't use it very much—not really enough to comment. On a personal note, during the last 12 months, I have been going through a process of appointments and getting my son diagnosed for autism and then ongoing therapies/appointments. I also have 2 other children and am expecting a third plus working part time so I have found adding this extra facet into my life almost impossible. It certainly has nothing against the resource. I have simply been too busy to put the time in and for that I apologize.* [Participant ID 51; accessed PCHMS: once]

The lack of importance participants placed on asthma:

> *Complacency—I should I know but having been asthmatic all my life I don't give it the importance I should.* [Participant ID 1; accessed PCHMS: once]
>
> *I guess I always think it will never get worse...which I know is wrong.* [Participant ID 47; accessed PCHMS: twice]

A belief that a written AAP is "irrelevant" to their condition:

> *Look, for someone who has just been diagnosed with asthma or someone quite young, it's probably great. But for someone like me who has had asthma for over 40 years, has an informal plan of what to do (ie, I*

> *know when I need to be on the preventative, what causes it, when I need Ventolin, what to do if I'm having too much Ventolin, etc), it's not very helpful.* [Participant ID 8; accessed PCHMS: twice]
>
> *I had one, but because my asthma triggers and symptoms and signs change so often they quickly become out of date.* [Participant ID 21; accessed PCHMS: twice]

Inadequacies of the intervention or asthma content:

> *Do I have to log in? It would be better if access was open.* [Participant ID 41; accessed PCHMS: 4 times]
>
> *I already understand my asthma. I thought this might contribute to that understanding, but I think it was aimed at a much younger/newer to asthma participant.* [Participant ID 8; accessed PCHMS: twice]

Competing Priorities on Health and Asthma

Participants in both groups were asked to report monthly on their life priorities and the importance they placed on their health and asthma. On a scale from 1 to 10 (where 1 was highest priority and 10 was lowest), participants on average rated health moderately highly (control: mean 3.5, SD 1.9; intervention: mean 3.3, SD 2.1). However, the priority they placed on asthma was not as high (control: mean 4.3, SD 2.2; intervention: mean 3.8, SD 2.3). In fact, asthma was often not a health issue reported by participants that caused them the most concern on a monthly basis.

The average number of life priorities reported by participants was similar in both groups (control: mean 3.0, SD 1.3; intervention: mean 2.7, SD 1.5). These priorities ranged from issues related to work, family/relationship, and money. Health was not always mentioned in this list of priorities. On average, participants reported approximately 2 health issues per month (control: mean 1.9, SD 1.0; intervention: mean 1.5, SD 0.9). These issues are related to a range of bodily systems (eg, cardiovascular, musculoskeletal, psychological, neurological) and not only restricted to the respiratory system.

Usage and Perception of Healthy.me

Participant usage and perceptions of the Healthy.me intervention are outlined in Tables 5 and 6. Most participants (80.5%, 124/154) did not access the intervention or accessed it only once (Table 5). Only one person accessed the intervention 10 times or more in this study (Table 5). Because only 30 participants used the website more than once, there was insufficient usage of the online features to make meaningful interpretation of their efficacy. On the TAM scale of 1 (strongly disagree) to 7 (strongly agree), participants indicated, on average, a neutral score (4.7-4.9) for the system's perceived usefulness and a neutral score for ease of use (4.8-4.9) (Table 6).

Table 5. Usage frequency of Healthy.me (n=154 participants).

Usage frequency (times)	Participants, n (%)
0	30 (19.5)
1	94 (61.0)
2-5	27 (17.5)
6-10	2 (1.3)
>10	1 (0.6)

Table 6. Perception of the intervention as measured by the Technology Acceptance Model (n=56).[a]

Perception of intervention	Mean (SD)
Perceived ease of use [b]	
Healthy.me was easy to use	4.9 (1.5)
I find it was easy to get Healthy.me to do what I wanted it to do	4.8 (1.4)
It was easy to become confident with using Healthy.me	4.9 (1.4)
Perceived usefulness [b]	
Managing my asthma through Healthy.me will be beneficial to me	4.7 (1.2)
The advantages of using Healthy.me to manage my asthma will outweigh the disadvantages	4.9 (1.2)
Overall, using Healthy.me will help me improve my asthma in general	4.8 (1.2)

[a] Participants were allocated to PCHMS and completed the poststudy questionnaire.

[b] Likert scale 1 to 7, where 1=strongly disagree, 4=neutral, 7=strongly agree.

Discussion

Principal Findings

Access to the Healthy.me PCHMS did not improve the rate of possession of written AAPs, planned visits to health services for nonemergency asthma management, asthma status, control, or work and study productivity. These results are in stark contrast to earlier trials of the same PCHMS, which showed significant improvements in outcomes associated with consumer behavior change, including influenza vaccination [7] and sexually transmitted infection screening [8]. The negative results also occur in a context where intervention and study designs were in accord with factors typically associated with successful uptake and efficacy of online consumer interventions.

High attrition rates are common in eHealth intervention studies [14], with a recent systematic review revealing that completion of protocol rates for depression sites ranged from 43% to 99% [34]. This study suffered from moderate to high rates of attrition in the intervention (64%) and control (45%) groups. A previous study of Healthy.me did not experience this degree of attrition [7]. However, that study was conducted in 2009, its duration was shorter, and the mean participant age was 26.2 years compared to 40 years among those allocated to the intervention in this RCT [7].

Utilization and Benefit

In decision theoretic terms, the expected utility of any eHealth intervention is a product of the utility or benefit of each individual interaction with the system to the user and the number of times the interaction takes place [35]. Systems that engage their users and, as a result, are used frequently are theoretically more likely to deliver benefit. This is reflected in the research evidence, where consumer eHealth systems seem to demonstrate a clear dose-response relationship between use and benefit [35].

In this trial, some participants suggested that they saw little benefit in using the system, either because they or their health professionals saw little value in having an AAP, because asthma management was not a major priority in their life compared to other competing priorities, or that they have already developed their own strategies to manage the condition and needed no further assistance. Perhaps as suggested by some participants, the intervention would be more helpful for those who are newly diagnosed with asthma.

Comparison With Prior Work

A systematic review of PHRs used for chronic conditions found that unless a system clearly assisted consumers in self-management tasks, they were unlikely to be successful [36]. This benefit might come from tracking important parameters to control an illness, such as blood pressure or glucose levels, or by delivering feedback when changes to management are needed. That review identified diabetes, hypertension, asthma, HIV, fertility management, glaucoma, and hyperlipidemia as having the most evidence for PHR benefit. However, only one study in that review actually included asthma patients and these were grouped together with other patients who had diabetes or hypertension [36]. The only outcome measure was patient activation, and asthma patients represented only 7% of the sample, providing weak evidence of PHR benefit in asthma management outcomes.

Recent systematic reviews concluded that although there is evidence that some digital interventions are associated with positive asthma self-management outcomes [37,38], most interventions do not use behavioral change theory, clinical guidelines, and/or assessment tools to inform their design [37]. A Cochrane review on smartphone apps for asthma concluded there is currently lack of evidence to advise clinical practitioners, policy makers, and the general public on ways to implement these interventions for asthma self-management programs [39]. Relevant to the AAP, a theoretical model has proposed 4 elements that are essential in facilitating the “right” contexts between patients and professionals, but few studies have used all these elements in their implementation [40].

Our own earlier trials of this intervention focused on supporting preventive health tasks. A trial aimed at encouraging influenza vaccination demonstrated a significant doubling in vaccination rates, most likely because the system allowed easy and immediate access to booking a vaccination with a primary care center, for a condition where seasonality and acting in a timely manner is important [7]. Similar benefits were demonstrated when the system was targeted at increasing screening rates among young adults for sexually transmitted infections, where use of an online booking system may have additionally reduced any sense of stigma associated with making a decision to act [8].

Lessons Learned

Although the lack of uptake of eHealth interventions is a widely known phenomenon [13-15], the literature on negative findings in this field is still scarce. Our study provides a number of lessons:

1. *Consumers must perceive the need for assistance with a task.* Even though the research evidence clearly demonstrates the value of an AAP, its low level of adoption in the population and the commentary received in this study suggests that at least some adults with asthma either do not agree or have yet to be convinced.
2. *Consumers must assign priority to the tasks supported by the intervention.* Participants in this study assigned a low priority to their asthma management compared with other life priorities. The mean age of participants in this study was 40 and most reported living with multiple competing priorities (eg, work commitment, lack of time) and other health concerns (eg, multimorbidity).
3. *The cost of adoption of the intervention must be lower than the benefit.* Our PCHMS was a stand-alone system that did not integrate into other apps participants might already have been using, such as diaries and social media. It consequently required additional effort to use. A substantial number of participants were recruited via a Facebook social network related to asthma and used Facebook several times a day. These individuals may have had higher expectations of the intervention regarding the degree of system integration, content, social network size, and the overall “polish” of the system.

4. *Outcome measures must be relevant to consumers and providers.* Although the primary outcome used in this RCT—possession of an up-to-date written AAP—is an indicator of recommended care for asthma [41], it is essential to consider how relevant and important the outcome measure is to *both* consumers and providers. Although there is evidence supporting the efficacy of the written AAP, there is possibly the misconception that it is only useful for those newly diagnosed with asthma. Perhaps there needs to be more emphasis on uncovering how relevant an outcome measure is to both consumers and providers before attempts are made to influence behavior change.

Implications for Consumer eHealth Design

Design for Attrition

Although current evidence advocates the importance of having a theoretical basis to direct behavioral changes, it is equally important to consider whether such theories can be used to minimize participant attrition. For example, identifying early on those who are truly uninterested and focus instead on those who are likely to continue could potentially reduce participant attrition [14]. Perhaps all interventions should be designed with a plan to minimize participant dropout before commencing participant recruitment.

Design for Implementation

Studies have confirmed once again that implementation uptake is often the biggest challenge in any eHealth project, both for consumers and clinicians. Trials that focus on implementation of asthma interventions are emerging in clinical settings [42,43]. Yet, implementation strategies that consider consumer settings, their comorbidities, and their competing demands are lacking. Understanding how these consumer factors affect the uptake of an intervention is important. A recent review on digital interventions for asthma concluded patient perspectives are often largely ignored [38]. Perhaps the next generation of digital intervention should incorporate consumer-clinician implementation strategies at the core of every digital intervention design.

Design for Context

Rather than attempting to “perfect” the design of an intervention to exist on its own, interventions should be designed for the context. When designing an intervention for consumers and patients, it is important to identify early on whether the intervention should focus on task support or on belief change. Moreover, research should focus on how we can design consumer eHealth interventions that are integrated in health care settings and/or how such interventions would function in the consumer circle of care (eg, caregivers).

Limitations

Study strengths include nationwide recruitment, use of recommended care indicators for outcome measures, and triangulation of participant feedback with quantitative results.

Notable limitations of this study include the gender and age distribution of participants, the attrition rate, and the use of self-reported data. The majority of participants were female in their late thirties / early forties and it is possible this population sample behaved differently than a more representative sample.

Participants had a higher rate of AAP possession than reported in other studies. As a result, as a cohort, they may already be better engaged and confident in their self-management and less likely to benefit from the intervention compared to the population average, reducing the potential effect size. Further, because the outcome measure was focused on having an up-to-date written AAP that was updated by a clinician (eg, once a year), we may have missed some participants as the study duration was only 12 months. Future studies should consider extending the trial period to more than 12 months.

Our primary recruitment strategy is online, which has a number of limitations, such as high rates of attrition. More effective recruitment could potentially result when it is channeled through influencers such as health care providers or with the encouragement of caregivers who help patients to deal with issues every day. However, this is an intervention designed primarily for consumers, to be delivered online, thus it is important that there is a direct channel to recruit consumers who are already online.

Conclusions

Consumers are increasingly turning to the Internet and social media for health advice, yet we still do not fully understand why some online interventions work and others do not. In this study, participant goals were poorly aligned with the clinical goals of the system despite there being clear evidence underpinning these latter clinical goals. It may be that a different approach is required in the domain of asthma management in adults, at least as far as AAPs are concerned, focusing not so much on task support as on belief change. More generally, researchers should not feel discouraged to publish negative findings because in failure many significant lessons can be learned.

Acknowledgments

The authors thank Sara Morgan from the Asthma Foundation New South Wales for her assistance in designing the intervention content and Nathan Mortimer for his assistance in forum moderation. The authors also thank Jingbo Liu, Vitaliy Kim, Farshid Anvari, and Jay Liu for their contributions to software development, and the study participants for their time and feedback. This research is supported by grants received from the National Health and Medical Research Council (NHMRC) Centre of Research Excellence in Informatics and EHealth (1032664). The funding body did not have a role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

Conflicts of Interest

The University of New South Wales, Enrico Coiera, and Annie Lau could benefit from the commercial exploitation of the Healthy.me platform or its technologies.

Multimedia Appendix 1

Use of Health Belief Model (HBM) in Asthma Patient Journeys.

[PDF File (Adobe PDF File), 4KB - jmir_v17i12e283_app1.pdf]

Multimedia Appendix 2

Utilization of Personally Controlled Health Management System (PCHMS).

[PDF File (Adobe PDF File), 47KB - jmir_v17i12e283_app2.pdf]

References

1. Institute of Medicine. Health IT and Patient Safety: Building Safer Systems for Better Care. Washington, DC: The National Academies Press; 2011.
2. Lau AYS, Parker A, Early J, Sacks G, Anvari F, Coiera E. Comparative usage of a web-based personally controlled health management system and normal support: a case study in IVF. electronic Journal of Health Informatics 2013;7(2):16.
3. Wagner PJ, Dias J, Howard S, Kintziger KW, Hudson MF, Seol Y, et al. Personal health records and hypertension control: a randomized trial. J Am Med Inform Assoc 2012 Aug;19(4):626-634 [FREE Full text] [doi: 10.1136/amiajnl-2011-000349] [Medline: 22234404]
4. Tang PC, Overhage JM, Chan AS, Brown NL, Aghighi B, Entwistle MP, et al. Online disease management of diabetes: engaging and motivating patients online with enhanced resources-diabetes (EMPOWER-D), a randomized controlled trial. J Am Med Inform Assoc 2013 May 1;20(3):526-534 [FREE Full text] [doi: 10.1136/amiajnl-2012-001263] [Medline: 23171659]
5. Grant RW, Wald JS, Schnipper JL, Gandhi TK, Poon EG, Orav EJ, et al. Practice-linked online personal health records for type 2 diabetes mellitus: a randomized controlled trial. Arch Intern Med 2008 Sep 8;168(16):1776-1782 [FREE Full text] [doi: 10.1001/archinte.168.16.1776] [Medline: 18779465]
6. Bourgeois FT, Simons WW, Olson K, Brownstein JS, Mandl KD. Evaluation of influenza prevention in the workplace using a personally controlled health record: randomized controlled trial. J Med Internet Res 2008;10(1):e5 [FREE Full text] [doi: 10.2196/jmir.984] [Medline: 18343794]
7. Lau Annie Y S, Sintchenko V, Crimmins J, Magrabi F, Gallego B, Coiera E. Impact of a web-based personally controlled health management system on influenza vaccination and health services utilization rates: a randomized controlled trial. J Am Med Inform Assoc 2012;19(5):719-727 [FREE Full text] [doi: 10.1136/amiajnl-2011-000433] [Medline: 22582203]
8. Mortimer NJ, Rhee J, Guy R, Hayen A, Lau AYS. A Web-based personally controlled health management system increases sexually transmitted infection screening rates in young people: a randomized controlled trial. J Am Med Inform Assoc 2015 Mar 15:805-814. [doi: 10.1093/jamia/ocu052] [Medline: 25773130]
9. Schnipper JL, Gandhi TK, Wald JS, Grant RW, Poon EG, Volk LA, et al. Effects of an online personal health record on medication accuracy and safety: a cluster-randomized trial. J Am Med Inform Assoc 2012;19(5):728-734 [FREE Full text] [doi: 10.1136/amiajnl-2011-000723] [Medline: 22556186]
10. Tiong S, Delaney G, Lau AYS, Koh E, Adams D, Bell V. Utilisation and acceptability of an e-Health strategy to facilitate care of breast cancer survivors. Asia-Pac J Clin Onco 2012 Nov;8:330.
11. Lau AYS, Proudfoot J, Andrews A, Liaw ST, Crimmins J, Arguel A, et al. Which bundles of features in a Web-based personally controlled health management system are associated with consumer help-seeking behaviors for physical and emotional well-being? J Med Internet Res 2013 May;15(5):e79 [FREE Full text] [doi: 10.2196/jmir.2414] [Medline: 23649790]
12. Lau AYS, Dunn AG, Mortimer N, Gallagher A, Proudfoot J, Andrews A, et al. Social and self-reflective use of a Web-based personally controlled health management system. J Med Internet Res 2013;15(9):e211 [FREE Full text] [doi: 10.2196/jmir.2682] [Medline: 24064188]
13. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006;8(3):e20; author reply e21 [FREE Full text] [doi: 10.2196/jmir.8.3.e20] [Medline: 17032636]
14. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
15. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. BMJ 2010;341:c5814 [FREE Full text] [Medline: 21081595]

16. Price M, Bellwood P, Kitson N, Davies I, Weber J, Lau F. Conditions potentially sensitive to a personal health record (PHR) intervention, a systematic review. BMC Med Inform Decis Mak 2015;15:32 [FREE Full text] [doi: 10.1186/s12911-015-0159-1] [Medline: 25927384]
17. Masoli M, Fabian D, Holt S, Beasley R. Global burden of asthma. Global Initiative for Asthma 2004:na.
18. Australia National Asthma Council. Asthma Management Handbook. Melbourne: National Asthma Council; 2006.
19. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev 2003(1):CD001117. [doi: 10.1002/14651858.CD001117] [Medline: 12535399]
20. Gibson PG, Powell H. Written action plans for asthma: an evidence-based review of the key components. Thorax 2004 Feb;59(2):94-99 [FREE Full text] [Medline: 14760143]
21. Gibson PG. Asthma action plans: use it or lose it. Prim Care Respir J 2004 Mar;13(1):17-18 [FREE Full text] [doi: 10.1016/j.pcrj.2003.12.001] [Medline: 16701632]
22. Australian Institute of Health and Welfare. Asthma in Australia 2011. Canberra: Australian Centre for Asthma Monitoring; 2011.
23. Arguel A, Lau AYS, Dennis S, Liaw S, Coiera E. An internet intervention to improve asthma management: rationale and protocol of a randomized controlled trial. JMIR Res Protoc 2013;2(2):e28 [FREE Full text] [doi: 10.2196/resprot.2695] [Medline: 23942523]
24. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
25. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [Medline: 6392204]
26. Bandura A. Human agency in social cognitive theory. Am Psychol 1989 Sep;44(9):1175-1184. [Medline: 2782727]
27. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
28. Little R, Yau L. Intent-to-treat analysis for longitudinal studies with drop-outs. Biometrics 1996 Dec;52(4):1324. [doi: 10.2307/2532847]
29. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: current practice and problems. Stat Med 2002 Oct 15;21(19):2917-2930. [doi: 10.1002/sim.1296] [Medline: 12325108]
30. Reddel HK, Taylor DR, Bateman ED, Boulet L, Boushey HA, Busse WW, American Thoracic Society/European Respiratory Society Task Force on Asthma ControlExacerbations. An official American Thoracic Society/European Respiratory Society statement: asthma control and exacerbations: standardizing endpoints for clinical asthma trials and clinical practice. Am J Respir Crit Care Med 2009 Jul 1;180(1):59-99. [doi: 10.1164/rccm.200801-060ST] [Medline: 19535666]
31. Juniper EF, Bousquet J, Abetz L, Bateman ED. Identifying 'well-controlled' and 'not well-controlled' asthma using the Asthma Control Questionnaire. Respir Med 2006 Apr;100(4):616-621 [FREE Full text] [doi: 10.1016/j.rmed.2005.08.012] [Medline: 16226443]
32. Braido F. Failure in asthma control: reasons and consequences. Scientifica 2013:1-15. [doi: 10.1155/2013/549252]
33. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
34. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
35. Coiera E. Guide to Health Informatics (3rd edition). London: CRC Press; 2015.
36. Solomon M, Wagner SL, Goes J. Effects of a Web-based intervention for adults with chronic conditions on patient activation: online randomized controlled trial. J Med Internet Res 2012;14(1):e32 [FREE Full text] [doi: 10.2196/jmir.1924] [Medline: 22353433]
37. Al-Durra M, Torio M, Cafazzo JA. The use of behavior change theory in Internet-based asthma self-management interventions: a systematic review. J Med Internet Res 2015;17(4):e89 [FREE Full text] [doi: 10.2196/jmir.4110] [Medline: 25835564]
38. Morrison D, Wyke S, Agur K, Cameron EJ, Docking RI, Mackenzie AM, et al. Digital asthma self-management interventions: a systematic review. J Med Internet Res 2014;16(2):e51 [FREE Full text] [doi: 10.2196/jmir.2814] [Medline: 24550161]
39. Marcano BJ, Huckvale K, Greenfield G, Car J, Gunn LH. Smartphone and tablet self management apps for asthma. Cochrane Database Syst Rev 2013;11:CD010013. [doi: 10.1002/14651858.CD010013.pub2] [Medline: 24282112]
40. Ring N, Jepson R, Pinnock H, Wilson C, Hoskins G, Wyke S, et al. Developing novel evidence-based interventions to promote asthma action plan use: a cross-study synthesis of evidence from randomised controlled trials and qualitative studies. Trials 2012;13:216 [FREE Full text] [doi: 10.1186/1745-6215-13-216] [Medline: 23164151]
41. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: assessing the appropriateness of health care delivery in Australia. Med J Aust 2012 Jul 16;197(2):100-105. [Medline: 22794056]

42. van Gaalen JL, Bakker MJ, van Bodegom-Vos BL, Snoeck-Stroband JB, Assendelft WJ, Kaptein AA, et al. Implementation strategies of internet-based asthma self-management support in usual care. Study protocol for the IMPASSE cluster randomized trial. Implement Sci 2012;7:113 [FREE Full text] [doi: 10.1186/1748-5908-7-113] [Medline: 23171672]
43. Mold JW, Fox C, Wisniewski A, Lipman PD, Krauss MR, Harris DR, et al. Implementing asthma guidelines using practice facilitation and local learning collaboratives: a randomized controlled trial. Ann Fam Med 2014;12(3):233-240 [FREE Full text] [doi: 10.1370/afm.1624] [Medline: 24821894]

Abbreviations

AAP: asthma action plan
ACQ: Asthma Control Questionnaire
HBM: Health Belief Model
LOCF: last observation carried forward
PCHMS: personally controlled health management system
PHR: personal health record
RCT: randomized controlled trial
TAM: Technology Assessment Model

Edited by G Eysenbach; submitted 21.05.15; peer-reviewed by A Chang, C Serrano, K Nazi, N Archer; comments to author 19.08.15; revised version received 21.10.15; accepted 09.11.15; published 15.12.15.

Please cite as:
Lau AYS, Arguel A, Dennis S, Liaw ST, Coiera E
"Why Didn't it Work?" Lessons From a Randomized Controlled Trial of a Web-based Personally Controlled Health Management System for Adults with Asthma
J Med Internet Res 2015;17(12):e283
URL: http://www.jmir.org/2015/12/e283/
doi:10.2196/jmir.4734
PMID:26678294

Original Paper

Large-Scale Survey Findings Inform Patients' Experiences in Using Secure Messaging to Engage in Patient-Provider Communication and Self-Care Management: A Quantitative Assessment

Jolie N Haun[1,2], PhD, EdS; Nitin R Patel[1], MPH; Jason D Lind[1], MPH, PhD; Nicole Antinori[1], MBA

[1]Veterans Health Administration, HSR&D Center of Innovation on Disability and Rehabilitation Research, Tampa, FL, United States
[2]Department of Community & Family Health, College of Public Health, University of South Florida, Tampa, FL, United States

Corresponding Author:
Jolie N Haun, PhD, EdS
Veterans Health Administration
HSR&D Center of Innovation on Disability and Rehabilitation Research
8900 Grand Oak Circle (118M)
Tampa, FL, 33637
United States
Phone: 1 904 377 6796
Fax: 1 813 558 3990
Email: joliehaun@gmail.com

Abstract

Background: Secure email messaging is part of a national transformation initiative in the United States to promote new models of care that support enhanced patient-provider communication. To date, only a limited number of large-scale studies have evaluated users' experiences in using secure email messaging.

Objective: To quantitatively assess veteran patients' experiences in using secure email messaging in a large patient sample.

Methods: A cross-sectional mail-delivered paper-and-pencil survey study was conducted with a sample of respondents identified as registered for the Veteran Health Administrations' Web-based patient portal (My HealtheVet) and opted to use secure messaging. The survey collected demographic data, assessed computer and health literacy, and secure messaging use. Analyses conducted on survey data include frequencies and proportions, chi-square tests, and one-way analysis of variance.

Results: The majority of respondents (N=819) reported using secure messaging 6 months or longer (n=499, 60.9%). They reported secure messaging to be helpful for completing medication refills (n=546, 66.7%), managing appointments (n=343, 41.9%), looking up test results (n=350, 42.7%), and asking health-related questions (n=340, 41.5%). Notably, some respondents reported using secure messaging to address sensitive health topics (n=67, 8.2%). Survey responses indicated that younger age (P=.039) and higher levels of education (P=.025) and income (P=.003) were associated with more frequent use of secure messaging. Females were more likely to report using secure messaging more often, compared with their male counterparts (P=.098). Minorities were more likely to report using secure messaging more often, at least once a month, compared with nonminorities (P=.086). Individuals with higher levels of health literacy reported more frequent use of secure messaging (P=.007), greater satisfaction (P=.002), and indicated that secure messaging is a useful (P=.002) and easy-to-use ($P \leq .001$) communication tool, compared with individuals with lower reported health literacy. Many respondents (n=328, 40.0%) reported that they would like to receive education and/or felt other veterans would benefit from education on how to access and use the electronic patient portal and secure messaging (n=652, 79.6%).

Conclusions: Survey findings validated qualitative findings found in previous research, such that veterans perceive secure email messaging as a useful tool for communicating with health care teams. To maximize sustained utilization of secure email messaging, marketing, education, skill building, and system modifications are needed. These findings can inform ongoing efforts to promote the sustained use of this electronic tool to support for patient-provider communication.

(J Med Internet Res 2015;17(12):e282) doi:10.2196/jmir.5152

KEYWORDS

cross-sectional survey; email; Internet communication tools; veterans

Introduction

Patient-provider communication is central in delivering high quality of care and promoting positive patient outcomes [1]. Electronic asynchronous secure email messaging within Web-based patient portals is gaining popularity as a viable efficient form of patient-provider communication [2-4]. Secure email messaging is a priority in the United States, as part of a national transformation initiative to create new models of care to support patient-provider communication and promote self-care management within the context of the patient-centered medical home [5].

Electronic communication, such as secure messaging, has been shown to be effective in supporting self-care management, patient engagement, and efficient use of health services [2-4,6-8]. A systematic review of literature suggested that obtained data moderately support the use of secure messaging, to improve glucose outcomes and increase patient satisfaction, and that secure messaging as part of an electronic patient portal is more effective than secure messaging alone [9]. Although not as strong, some reports also suggest that there is some evidence that adding a Web-based pharmacist to secure messaging improves blood pressure outcomes in patients with hypertension, and that secure messaging within an electronic portal improves ulcerative colitis symptoms and adherence to colorectal cancer screenings or heart failure management [9]. Although some studies have seen a positive effect on utilization, it should be noted that the systematic review did find evidence that secure messaging may positively or negatively affect efficiency or utilization [9].

Recognizing that the implementation of secure messaging is widespread and is quickly becoming a common practice as part of services provided by integrated electronic patient portal services, efforts to understand patients' experiences and needs when using secure messaging tools are warranted. This approach is imperative to supporting patients' sustained use of electronic communication mechanisms, such as secure messaging.

Consistent with this consumer-centric approach, our previous qualitative research, consistent with other published studies, indicates patients' value secure messaging as an efficient means of communication with their providers [10]. Benefits reported in a previous qualitative study included 24-hour access, avoiding phone calls and travel to health care facilities, and in general, saving time [10]. Our previous research provided insights into patients' experiences, but larger quantitative data studies are needed to understand veterans' experiences in using secure messaging and determine convergence with previous qualitative findings.

This paper presents findings of a cross-sectional survey study with a large sample of veterans who opted-in to use secure messaging to assess their reported experiences in using secure messaging and to evaluate factors that predict use and perceptions associated with using secure messaging. Findings from this survey research will inform efforts to quantify (1) veterans' reported use of the secure messaging tool, (2) their reasons for using secure messaging, and (3) factors that influence their use of secure messaging. Gathering these data in a large representative sample can inform efforts to develop education and marketing content for potential users, identify points of intervention to support sustained secure messaging use, and continue the accumulation of reported evidence on the use of electronic forms of patient-provider communication.

Methods

Study Design Overview

This is a cross-sectional study. A paper-and-pencil survey via mail was sent to veterans who had registered for the Veteran Health Administrations' Web-based patient portal (My HealtheVet) and opted to use secure messaging. The survey collected demographic data, assessed computer and health literacy, and secure messaging use.

Setting and Participants

The two-site study was conducted at 2 large Department of Veterans Affairs (VA) Medical Centers (VAMCs): the James A. Haley Veterans' Hospital (Tampa, FL, USA) and the Veterans Affairs Boston Healthcare System (Boston, MA, USA). We used administrative data to identify veterans at both VAMCs who had registered for My HealtheVet, completed the in-person process of authenticating their identity, and accessed the system to "opt-in" to use secure messaging. We then used randomization to create contact lists of 2100 potential participants. Of the 2100, 2073 (1022 in Boston; 1051 in Tampa) had complete information to mail a survey for completion. Veterans received US $10 for completing the survey. This study was approved and regulated by the VA Central Institutional Review Board.

Data Collection Instruments

The survey collected demographic data, assessed health literacy and eHealth literacy, and secure messaging use and perceptions. The majority of items were developed based on qualitative findings from a previous qualitative study conducted by the research team. Validated measures included the BRIEF Health Literacy Screening Tool [11,12], Computer-Email-Web (CEW) Fluency Scale [13], and The eHealth Literacy Scale (eHEALS) [14].

BRIEF Health Literacy Screening Tool

This is a 4-item screening tool to assess health literacy skills with 5-point Likert-type scale response options [11,12]. Score range was 4-20; score levels were as follows: 4-12=inadequate, 13-16=marginal, and 17-20=adequate. The correlation results of Rapid Estimate of Adult Literacy in Medicine (REALM) and Test of Functional Health Literacy in Adults-Short Form (STOFHLA) were as follows: $r=.40$, $P<.01$ for the BRIEF/REALM; $r=.42$, $P<.01$ for the BRIEF/STOFHLA; and $r=.61$, $P<.01$ for the REALM/STOFHLA. A principal component analysis suggested that the BRIEF measures one distinct construct "health literacy" (eigenvalue = 2.388), which accounted for 60% of score variance.

Computer-Email-Web Fluency Scale

The CEW Fluency Scale is an 18-item measure of common computer skills, with 5-point Likert-type scale response options (eg, not at all; very well) with a score range of 18-90. Cronbach alphas were established for subscales including computer fluency (alpha=.72), email fluency (alpha=.75), Web navigation (alpha=.64), and Web editing (alpha=.79) [13].

eHealth Literacy Scale

The eHEALS is a 10-item measure of eHealth literacy developed to measure consumers' knowledge, comfort, and perceived skills at finding, evaluating, and applying electronic health information to health problems. eHEALS items have 5-point Likert-type scale response options, with a possible score range of 10-50. Previous validation results suggested internal consistency reliability to be .88, and test-retest reliability (*r*) from baseline to 6-month follow-up to be .68-.40. Principal components analysis produced a single factor (56% of variance) [14].

Data Analysis

All survey data in this study were stored on a secure VA network. Analyses were managed using the statistical software suite SPSS (SPSS Inc, Chicago, IL, USA). Frequencies and proportions were computed for categorical variables and mean and standard deviations were computed for continuous variables to describe sample characteristics and provide a descriptive overview of survey findings. Chi-square tests were conducted to assess association between categorical variables and one-way analysis of variance was conducted to assess significant differences in means for continuous variables.

Results

Participants

Of the 2073 surveys mailed (1022 in Boston; 1051 in Tampa), 819 respondents provided completed survey data for analysis.

Survey Findings

The majority of participants were older, white, non-Hispanic/non-Latino males, with an average age of 62 years (data not shown). Most participants had at least a high-school education and more than half (n=434, 53.0%) had an annual income of US $35,001 or more. Demographic data are presented in Table 1.

Most survey respondents reported everyday use of computer (n=662, 81.0%) and Internet (n=653, 79.7%). The majority of respondents reported using My HealtheVet a "few times a month or less" (n=629, 76.8%); and using secure messaging 6 months or longer (n=499, 60.9%). Most participants (486, 59.3%) reported using secure messaging at least once a year, and 131 (16.0%) reported using it at least once a month. Tables 2-5 present data on patients' computer, Internet, My HealtheVet, and secure messaging use.

Pearson chi-square tests on demographic variables and secure messaging use indicate that younger age (*P*=.039) and higher levels of education (*P*=.025) and income (*P*=.003) are associated with more frequent use. Women were more likely to report using secure messaging more often (*P*=.098), but this only had a marginal significance. Minorities were more likely to report using secure messaging more often, at least once a month (*P*=.086). These findings and other demographic factors not significantly associated with secure messaging use are presented in Table 6.

Overall, respondents' views on secure messaging are as follows: a good communication tool (n=619, 75.6%); saves time (n=590, 72.0%); and easy to use (n=544, 66.4%). Although 689 (84.1%) respondents reported intention to use secure messaging in the future, 342 (41.8%) reported that secure messaging could be improved to make it a more useful tool. Many respondents (n=328, 40.0%) reported that they would like to receive education and/or support on how to use My HealtheVet and secure messaging to manage their health care. A vast majority of respondents felt that other veterans would benefit from education on how to access and use My HealtheVet and secure messaging (n=652, 79.6%). The vast majority of respondents reported secure messaging as being a safe and secure form of communication (n=585, 71.4%). Patient-reported experiences in using secure messaging are illustrated in Table 7.

Respondents reported that secure messaging is useful for completing medication refills (n=546, 66.7%), medication questions (n=313, 38.2%), managing appointments (n=343, 41.9%), test results (n=350, 42.7%), and health-related questions (n=340, 41.5%). Notably, a small percentage of respondents reported using secure messaging to address sensitive health topics (n=67, 8.2%). Reasons why patients find it helpful and reasons for its use are presented in Table 8.

Secure messaging usefulness and reasons for its use scores were significantly higher for veterans who reported more frequent use of computer (*P*<.002 and *P*<.012, respectively) and Internet (*P*<.001 and *P*<.001, respectively). Similarly, scores were significantly higher for veterans who reported using My HealtheVet at least once a week (*P*<.001 and *P*<.001, respectively) and using secure messaging at least once a month (*P*<.001 and *P*<.001, respectively). Those who reported using secure messaging for more than a year also had significantly higher scores than those reporting its use for shorter periods. Findings related to respondents' perceptions of usefulness and reasons for use by technology-use factors are presented in Table 9.

The BRIEF scores indicate that the majority of the sample had adequate health literacy (n=566, 69.1%); 174 (21.2%) had marginal health literacy and 77 (9.4%) had inadequate health literacy. The mean eHEALS and CEW Fluency Scale scores were 38.2 (SD 7.1; range 10-50) and 77.5 (SD 16.3; range 18-90), respectively. Respondents reporting higher levels of health literacy reported more frequent use of computers and the Internet (*P*≤.001), more frequent use of secure messaging (*P*=.007), and greater satisfaction with secure messaging (*P*=.002); additionally, they were more likely to report that it was a useful communication tool (*P*=.002), easy to use (*P*≤.001), and it as a safe and secure form of communication (*P*=.019). Interestingly, there were no differences in health literacy level based on reported intention to use secure email messaging in the future (*P*=.545). Those with lower levels of health literacy were more likely to request education and/or support (*P*≤.001). Individuals with higher eHEALS and CEW scores were also more likely to report more frequent use of computer, Internet, and secure email messaging (*P*≤.001 and *P*≤.001); greater satisfaction with the tool (*P*≤.001 and *P*=.013); they were also more likely to report that the tool was easy to use (*P*≤.001 and *P*≤.001), saves time (*P*≤.001 and *P*=.031), and is a safe and secure form of communication (*P*≤.001 and *P*≤.001). Individuals with higher eHEALS scores were also more likely to report intention to use secure email messaging in the future (*P*≤.001) and that secure email messaging was a useful communication tool (*P*≤.001). Statistics presenting relationships between respondents' eHEALS, CEW, and BRIEF scores and secure messaging use and satisfaction are presented in Table 10.

Table 1. Descriptive statistics on patient demographics (N=819).

Variable		n (%)
Study site		
	Boston	339 (41.4)
	Tampa	480 (58.6)
Gender		
	Male	711 (86.8)
	Female	107 (13.1)
	Missing	1 (0.1)
Minority status		
	Minority	93 (11.4)
	Nonminority	726 (88.6)
Ethnicity		
	Hispanic or Latino	45 (5.5)
	Not Hispanic or Latino	706 (86.2)
	Missing	68 (8.3)
Education		
	High school or less	126 (15.4)
	Some college/vocational school/associate degree	380 (46.4)
	Bachelor's degree	198 (24.2)
	Graduate degree	113 (13.8)
	Missing	2 (0.2)
Income		
	≤US $15,000 per year	101 (12.3)
	US $15,001-US $25,000 per year	129 (15.8)
	US $25,001-US $35,000 per year	138 (16.8)
	US $35,001-US $45,000 per year	132 (16.1)
	>US $45,000 per year	302 (36.9)
	Missing	17 (2.1)
Marital status		
	Not married	333 (40.7)
	Married	483 (59.0)
	Missing	3 (0.4)

Table 2. Patients' computer, Internet, and My HealtheVet use (N=819).

Response options	Never n (%)	Few times a month or less n (%)	At least once a week n (%)	Everyday n (%)
How often do you use a computer?	15 (1.8)	31 (3.8)	109 (13.3)	662 (80.8)
How often do you use the Internet?	15 (1.8)	33 (4.0)	116 (14.2)	653 (79.7)
How often do you use the My HealtheVet website?	37 (4.5)	629 (76.8)	136 (16.6)	10 (1.2)

Table 3. Data collection on participants' My HealtheVet use (N=819).

Response options	Yes n (%)	No n (%)	I do not know n (%)
Have you completed the in-person authentication to upgrade your My HealtheVet account to use tools such as the secure messaging feature and Blue Button?	656 (80.1)	68 (8.3)	89 (10.9)
Have you opted-in to use the secure messaging feature on My HealtheVet?	658 (80.3)	70 (8.5)	86 (10.5)
Have you used the secure messaging (VA's secure email) feature on My HealtheVet?	565 (69.0)	204 (24.9)	45 (5.5)

Table 4. Participants' secure messaging usage in general (N=819).

Response options	<6 months n (%)	6 months to 1 year n (%)	> 1 year n (%)	Does not apply n (%)
How long have you been using secure messaging?	171 (20.9)	187 (22.8)	312 (38.1)	133 (16.2)

Table 5. Participants' secure messaging usage in My HealtheVet (N=819).

Response options	Never n (%)	At least once a year n (%)	At least once a month n (%)	Does not apply n (%)
How often do you use secure messaging on the My HealtheVet website?	116 (14.2)	486 (59.3)	131 (16.0)	76 (9.3)

Discussion

Principal Results

Findings from this survey research provided data on (1) veterans' reported use of the secure messaging tool; (2) veterans' reported reasons for using secure messaging; and (3) factors that influence their use of secure messaging. Key findings from this cross-sectional survey suggest that in a random sample (N=819) of patients receiving care within the VA who opted in to use secure messaging, a majority reported using secure messaging at least once a year, with less than 15% reporting never using the communication tool. Our VA sample had a higher percentage of participants reporting use of secure messaging than previous survey studies outside the VA, in which about 10-37% of respondents reported using email to contact their physician [15-17]. However, our percentage is likely higher due to our sampling methods that identified veteran patients who opted in to use secure messaging. It is safe to assume that this percentage would decrease if the survey were completed by the general patient population.

Overall, respondents reported being satisfied with secure messaging, as it provides a safe and secure communication tool that is easy to use and saves time. These results are consistent with our previous research and other reports [4,10,16,18,19]. Respondents reported that secure messaging is useful for tasks such as completing medication refills, managing appointments, receiving test results, and addressing health-related questions. A key finding in this study is that a small percentage of respondents reported using secure messaging to address sensitive health topics. This suggests that secure messaging offers patients a confidential, secure, and safe space to bring up sensitive topics, such as erectile dysfunction and sexually transmitted diseases, and avoiding the stigma or embarrassment of discussing these topics in person.

Some research suggests that patient concerns about data security may prevent the uptake of electronic health records; however, a majority of our respondents felt that secure messaging is a safe and secure form of communication [4,10,11].

Our findings are consistent with a previously published research outside the VA, which found that older age was negatively associated with frequency of use to contact health care providers using email (ie, secure messaging) [16]. Consistent with previous findings, income was positively correlated with preferences to use email (ie, secure messaging) to communicate with health care providers [15]. Previous research has shown mixed findings about minority groups' preferences for using email-type services to communicate with health care providers [15,16]. However, in our sample, minority status was consistent with findings that suggest a positive association between minority status and use of electronic email communication with health care providers [16]. Further research is needed to better understand minority preferences and reasons for using electronic communication with health care providers.

Table 6. Pearson chi-square test between demographic variables and secure messaging use (N=819).

Demographic variable		How often do you use secure messaging?			
		Never	At least once a year	At least once a month	*P*-value
Age, mean (SD)		64.2 (13.1)	62.0 (12.9)	60.0 (13.1)	.039[a]
Study site, n (%)					
	Boston	56 (6.8)	190 (23.2)	55 (6.7)	.190
	Tampa	60 (7.3)	296 (36.1)	76 (9.3)	
Gender, n (%)					
	Male	106 (12.9)	415 (50.1)	109 (13.3)	.098[b]
	Female	9 (1.1)	71 (8.7)	22 (2.7)	
White/Caucasian, n (%)					
	No	14 (1.7)	48 (5.9)	22 (2.7)	.086[b]
	Yes	102 (12.5)	438 (53.5)	109 (13.3)	
Ethnicity, n (%)					
	Hispanic or Latino	7 (0.9)	24 (2.9)	8 (1.0)	.722
	Not Hispanic or Latino	97 (11.8)	431 (52.6)	108 (13.2)	
Education, n (%)					
	High school or less	28 (3.4)	66 (8.1)	25 (3.1)	.025[a]
	Some college/vocational School/associate degree	53 (6.5)	217 (26.5)	66 (8.1)	
	Bachelor's degree	23 (2.8)	123 (15.0)	26 (3.2)	
	Graduate degree	11 (1.3)	79 (9.6)	14 (1.7)	
Income, n (%)					
	≤US $15,000 per year	17 (2.1)	49 (6.0)	28 (3.4)	.003[a]
	US $15,001-US $25,000 per year	16 (2.0)	78 (9.5)	21 (2.6)	
	US $25,001-US $35,000 per year	29 (3.5)	79 (9.6)	19 (2.3)	
	US $35,001-US $45,000 per year	13 (1.6)	81 (9.9)	25 (3.1)	
	>US $45,000 per year	36 (4.4)	190 (23.2)	36 (4.4)	
Marital status, n (%)					
	Not married	48 (5.9)	203 (24.8)	60 (7.3)	.715
	Married	67 (8.2)	281 (34.3)	71 (8.7)	

[a] Significant at the .05 level.

[b] Significant at the .10 level.

Table 7. Patients' experience in using secure messaging (N=819).

Experience	Disagree n (%)	Neutral n (%)	Agree n (%)	Do not know n (%)
I am satisfied with the secure messaging feature on My HealtheVet.	43 (5.3)	84 (10.3)	572 (69.8)	107 (13.1)
I get responses to my secure messages in a timely fashion.	51 (6.2)	84 (10.3)	515 (62.9)	154 (18.8)
Secure messaging is a useful tool to communicate with health care providers.	23 (2.8)	48 (5.9)	619 (75.6)	117 (14.3)
Secure messaging is easy to use.	61 (7.4)	88 (10.7)	544 (66.4)	112 (13.7)
Secure messaging saves patients' time (eg, avoiding phone calls, and clinical visits).	30 (3.7)	70 (8.5)	590 (72.0)	115 (14.0)
Secure messaging could be improved to make it more useful to veterans.	62 (7.6)	215 (26.3)	342 (41.8)	186 (22.7)
Secure messaging is a secure and safe form of communication with VA providers.	18 (2.2)	71 (8.7)	585 (71.4)	134 (16.4)
I intend to use secure messaging in the future.	15 (1.8)	44 (5.4)	689 (84.1)	61 (7.4)
I would like to receive education and/or support on how to best use My HealtheVet and secure messaging to manage my health care.	185 (22.6)	246 (30.0)	328 (40.0)	49 (6.0)
Veterans would benefit from education on how to access and use My HealtheVet and secure messaging.	12 (1.5)	77 (9.4)	652 (79.6)	66 (8.1)

Table 8. Reasons patients find secure messaging helpful and reasons for use (N=819).

Reasons	Find secure messaging useful n (%)	Reason for using secure messaging n (%)
Medication refills	546 (66.7)	475 (58.0)
Medication questions	313 (38.2)	305 (37.2)
To manage appointments (eg, schedule, cancel)	343 (41.9)	301 (36.8)
Test results	350 (42.7)	292 (35.7)
Requests for tests	168 (20.5)	136 (16.6)
Request consult with specialist (eg, referral)	220 (26.9)	203 (24.8)
Health-related questions	340 (41.5)	301 (36.8)
Sensitive health topics (eg, sexually transmitted infections, mental health)	67 (8.2)	66 (8.1)
Can contact providers on my own time	425 (51.9)	381 (46.5)
Saves time compared with other method of communication (eg, phone)	448 (54.7)	377 (46.0)

In our study, respondents reporting higher levels of health literacy (BRIEF) and eHealth literacy (eHEALS and CEW) reported more frequent use of secure email messaging and greater satisfaction with secure messaging; besides, they were more likely to report that it was a safe, secure, and a useful communication tool. These findings are consistent with existing literature, suggesting that eHealth users tend to have higher levels of eHealth and health literacy [13,14,16,17,20].

Individuals with higher eHEALS were also more likely to report an intention to use secure email messaging in the future. Individuals with lower levels of health literacy were more likely to report a need for more education and/or support. Screening patients for their health literacy and eHealth literacy level may be an effective way to identify veterans with greater educational needs, and to allocate resources to support their use of tools such as secure messaging.

Table 9. Respondents' perceptions of usefulness and reasons for use by technology-use factors are presented (N=819).

	Reasons for usefulness of secure messaging			Reasons for which messaging is used		
	n	Mean (SD)	*P*	n	Mean (SD)	*P*
Frequency of computer use			.002			.012
Never	15	2.67 (1.72)		15	2.2 (1.66)	
Few times a month or less	31	2.74 (2.68)		31	1.9 (2.39)	
At least once a week	109	3.64 (2.75)		109	2.61 (2.40)	
Everyday	662	4.17 (2.74)		662	3.08 (2.45)	
Frequency of Internet use			.001			.001
Never	15	2.33 (1.72)		15	1.87 (1.69)	
Few times a month or less	33	2.82 (2.92)		33	1.82 (2.37)	
At least once a week	116	3.61 (2.67)		116	2.57 (2.32)	
Everyday	653	4.19 (2.73)		653	3.1 (2.46)	
Frequency of My HealtheVet use			<.001			<.001
Never	37	1.05 (2.4)		37	0.49 (1.37)	
Few times a month or less	629	3.9 (2.63)		629	2.82 (2.32)	
At least once a week	136	5.34 (2.51)		136	4.21 (2.45)	
Everyday	10	4.3 (3.59)		10	3.7 (3.59)	
Frequency of secure messaging use			<.001			<.001
Never	116	1.81 (2.47)		116	0.72 (1.51)	
At least once a year	486	4.5 (2.35)		486	3.43 (2.13)	
At least once a month	131	5.56 (2.36)		131	4.63 (2.34)	
Length of secure messaging use			<.001			<.001
<6 months	171	3.82 (2.45)		171	2.75 (2.28)	
6 months to 1 year	187	4.47 (2.4)		187	3.44 (2.27)	
>1 year	312	4.97 (2.5)		312	3.84 (2.27)	

A vast majority (80%) of respondents felt that other veterans would benefit from education on how to access and use My HealtheVet and secure messaging. Furthermore, data suggest respondents' perceptions of the usefulness of secure messaging are associated with frequency of use. These data warrant consideration for marketing secure messaging and providing education to intended users to ensure audiences understand the benefits and purposes for using this electronic communication tool. Finally, though the vast majority of participants were satisfied with the tool and reported intention to use secure messaging in the future, more than 40% reported that secure messaging tool could be improved to make it even more useful. This finding is timely and should be strongly considered as the VA continues efforts in redesigning and enhancing available electronic resources for their patients to support sustained use.

Table 10. Relationship between survey respondents' eHEALS, CEW, and BRIEF scores and My HealtheVet and secure messaging use.

	eHealth literacy score			Computer-Email-Web Fluency score			BRIEF Health Literacy score		
	n	Mean (SD)	*P*	n	mean (SD)	*P*	n	mean (SD)	*P*
Frequency of computer use			<.001			<.001			<.001
Never/few times per month or less	46	29.7 (7.5)		46	47.5 (24.1)		46	13.2 (4.6)	
At least once a week	108	34.8 (7.3)		108	67.4 (17.1)		109	16.3 (3.1)	
Everyday	661	39.4 (6.5)		661	81.2 (12.2)		662	17.8 (2.9)	
Frequency of Internet use			<.001			<.001			<.001
Never/few times per month or less	48	29.3 (7.3)		48	47.9 (23.6)		48	13.4 (4.4)	
At least once a week	115	34.6 (7)		115	67.1 (16.5)		116	16.2 (3.3)	
Everyday	652	39.5 (6.4)		652	81.5 (12.1)		653	17.9 (2.8)	
Frequency of My HealtheVet website use			<.001			.023			.998
Never/few times per month or less	664	37.8 (7.3)		664	77 (16.6)		664	17.4 (3.2)	
At least once a week/everyday	146	40.6 (5.8)		146	80.4 (13.3)		146	17.4 (3.3)	
Frequency of secure messaging use			<.001			<.001			<.007
Never	116	34.8 (8.3)		115	70.2 (20.2)		116	16.8 (3.3)	
At least once a year	484	38.9 (6.5)		486	79.3 (14.6)		484	17.6 (3.1)	
At least once a month	131	40 (6.1)		131	78.9 (14.5)		131	17 (3.5)	
Satisfied with secure messaging tool			<.001			.013			.002
Disagree	43	36.1 (7.6)		43	73.1 (20.5)		43	16.5 (3.7)	
Neutral	84	35.5 (7.1)		84	76.2 (15.8)		84	16.5 (3.7)	
Agree	570	39.3 (6.4)		571	79.2 (14.5)		570	17.6 (3.1)	
Secure messages receive response in a timely fashion			<.001			.001			<.001
Disagree	51	36.5 (7.5)		51	73.2 (16.4)		51	16.1 (3.4)	
Neutral	84	35.6 (7.6)		84	75.9 (15.7)		84	16.7 (3.7)	
Agree	513	39.7 (6.3)		515	80.1 (13.9)		513	17.8 (3.1)	
Secure messaging is a useful communication tool			<.001			.092			.002
Disagree	23	34.7 (8.9)		23	76.2 (17.4)		23	16.4 (3.9)	
Neutral	48	35.3 (5.9)		48	74.3 (15.8)		48	16.1 (4.2)	
Agree	617	39.2 (6.6)		618	79 (15)		617	17.6 (3.1)	
Secure messaging is easy to use			<.001			<.001			<.001
Disagree	61	34.7 (6.8)		61	70.9 (20.8)		61	15.9 (4.2)	
Neutral	87	35.3 (6.7)		88	73.6 (16.2)		88	17.2 (2.9)	
Agree	543	39.7 (6.3)		544	80.3 (13.6)		542	17.6 (3.1)	
Secure messaging saves time			<.001			.031			.084
Disagree	30	35.2 (7.9)		30	75.4 (14.5)		30	16.4 (3.8)	

	eHealth literacy score			Computer-Email-Web Fluency score			BRIEF Health Literacy score		
	n	Mean (SD)	*P*	n	mean (SD)	*P*	n	mean (SD)	*P*
Neutral	70	36 (6.8)		70	74.6 (17.8)		70	17 (3.2)	
Agree	588	39.2 (6.6)		589	79.1 (14.7)		588	17.5 (3.2)	
Secure messaging could be improved to make it more useful			.001			.014			.031
Disagree	62	41.4 (7)		62	83.5 (9.9)		62	18 (3.2)	
Neutral	215	38.9 (6.5)		215	78.1 (15.9)		215	17.5 (3)	
Agree	341	37.9 (7.1)		342	77.3 (16)		340	17 (3.4)	
Secure messaging is a secure and safe form of communication			<.001			.001			.019
Disagree	18	37.4 (8.5)		18	77.6 (16.2)		18	16.3 (3.8)	
Neutral	70	33.3 (8.3)		71	71.9 (18.6)		71	16.5 (3.8)	
Agree	584	39.4 (6.2)		585	79.1 (14.8)		584	17.5 (3.2)	
Intention to use secure messaging in the future			<.001			.059			.545
Disagree	15	34.1 (9.8)		15	73.9 (18.5)		15	17.4 (3.1)	
Neutral	44	34.5 (9.1)		44	73.2 (19.9)		44	16.8 (4.2)	
Agree	687	38.8 (6.6)		688	78.4 (15.3)		687	17.4 (3.2)	
Education and/or support on how to best use My HealtheVet and secure messaging would be helpful			<.001			<.001			<.001
Disagree	184	41.2 (7)		183	83.4 (10.8)		185	18.3 (3)	
Neutral	245	38.8 (6.9)		246	79.3 (14.7)		246	17.5 (3.2)	
Agree	328	36.2 (6.9)		328	73.2 (18)		326	16.8 (3.2)	
Veterans would benefit from education on how to access and use My HealtheVet and secure messaging			.372			.027			.022
Disagree	12	37.3 (11)		12	84.7 (9.2)		12	15.5 (4.8)	
Neutral	77	39.3 (6.5)		77	81.2 (12.8)		77	18 (3.3)	
Agree	651	38.1 (7.2)		651	76.9 (16.6)		651	17.2 (3.2)	

Limitations

The limitations of this cross-sectional survey study should be considered when interpreting these data. First, the generalizability of the survey sample in our study is a strength and a limitation. Our data are representative of the veteran patient population who are registered and opted-in to use secure messaging; however, these data do not represent those veterans who did not opt in to use this communication tool nor represent the general population's use of secure messaging systems outside of the VA. Furthermore, the response rate of our survey was less than 50%. Although consistent with response rates in similar user experience studies, caution should be exercised when generalizing these survey results to any veteran population. Second, our respondents were also more likely to be older white males, with higher levels of income and education. Although this is representative of the current veteran population, it is not representative of the diversification seen in younger active military and new veteran populations. Thus, it is best to exercise caution in generalizing our results to the entire veteran population; however, we can still draw useful conclusions from the survey data to understand veterans' experiences and reasons for use of secure messaging to inform future research in evaluating and increasing the sustained meaningful use of secure messaging. Third, as with any cross-sectional study, this survey does not allow statements on the causality of secure messaging use, however, it does provide much needed descriptive data to understand veteran's experiences in using secure messaging to manage their health. Finally, although this cross-sectional survey study provided important data on veterans' experiences and use of secure messaging, we cannot comment on how clinicians and other VA health care team members are using secure messaging to reciprocate communication with their patients or their experiences in using this tool. These questions require further examination.

Conclusions

Findings from this survey research provided data on veterans' reported use of the secure messaging tool, their reasons for using secure messaging, and factors that influence their use of secure messaging. These large-scale survey findings validated previously published qualitative findings suggesting that veterans perceive secure messaging as a useful tool for communicating with health care teams. Secure messaging use, perceptions of ease of use, and satisfaction differ by gender, education, income, health, and eHealth literacy levels. These data contribute to the body of knowledge on the use of electronic forms of patient-provider communication such as secure messaging and can be used to inform efforts to develop education and marketing content for potential users, as well as identify points of intervention to support sustained secure messaging use.

Acknowledgments

This research was supported by the VA, Veterans Health Administration, Office of Research and Development, Health Services Research and Development Service (HFP-09-156). This work was also supported in part by the Center of Innovation on Disability and Rehabilitation Research at the James A. Haley Veterans Hospital. The contents of this manuscript do not represent the views of the VA or the United States Government. The corresponding author, as a US government employee, has the right to grant on behalf of all authors, and does grant on behalf of all authors, a nonexclusive worldwide license to the publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to (1) publish, reproduce, distribute, display and store the contribution, (2) translate the contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the contribution, (3) create any other derivative work(s) based on the contribution, (4) to exploit all subsidiary rights in the contribution, (5) the inclusion of electronic links from the contribution to third party material where-ever it may be located; and, (6) licensee any third party to do any or all of the above.

Conflicts of Interest

None declared.

References

1. Committee on Quality of Health Care in America, Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century. Washington, DC: National Academy Press; 2001.
2. Andreassen HK, Trondsen M, Kummervold PE, Gammon D, Hjortdahl P. Patients who use e-mediated communication with their doctor: New constructions of trust in the patient-doctor relationship. Qual Health Res 2006 Feb;16(2):238-248. [doi: 10.1177/1049732305284667] [Medline: 16394212]
3. Roter DL, Larson S, Sands DZ, Ford DE, Houston T. Can e-mail messages between patients and physicians be patient-centered? Health Commun 2008;23(1):80-86. [doi: 10.1080/10410230701807295] [Medline: 18443995]
4. Houston TK, Sands DZ, Jenckes MW, Ford DE. Experiences of patients who were early adopters of electronic communication with their physician: Satisfaction, benefits, and concerns. Am J Manag Care 2004 Sep;10(9):601-608 [FREE Full text] [Medline: 15515992]
5. Office of the National Coordinator for Health Information Technology. New Stage 2 Core and Menu Objectives for EPs Providers & Professionals, HealthIT.gov. URL: http://www.healthit.gov/providers-professionals/achieve-meaningful-use/core-measures-2/use-secure-electronic-messaging [accessed 2013-09-23] [WebCite Cache ID 6Jr4yzCn1]
6. Zhou YY, Garrido T, Chin HL, Wiesenthal AM, Liang LL. Patient access to an electronic health record with secure messaging: Impact on primary care utilization. Am J Manag Care 2007 Jul;13(7):418-424 [FREE Full text] [Medline: 17620037]
7. Zhou YY, Kanter MH, Wang JJ, Garrido T. Improved quality at Kaiser Permanente through e-mail between physicians and patients. Health Aff (Millwood) 2010 Jul;29(7):1370-1375 [FREE Full text] [doi: 10.1377/hlthaff.2010.0048] [Medline: 20606190]
8. Harris LT, Haneuse SJ, Martin DP, Ralston JD. Diabetes quality of care and outpatient utilization associated with electronic patient-provider messaging: A cross-sectional analysis. Diabetes Care 2009 Jul;32(7):1182-1187 [FREE Full text] [doi: 10.2337/dc08-1771] [Medline: 19366959]
9. Goldzweig CL, Towfigh AA, Paige NM, Orshansky G, Haggstrom D, Beroes JEA. Systematic Review: Secure Messaging Between Providers and Patients, and Patients' Access to Their Own Medical Record Internet. Department of Veterans Affairs (US): Washington, DC; 2012. URL: http://www.ncbi.nlm.nih.gov/pubmedhealth/PMH0048536/ [accessed 2015-11-19] [WebCite Cache ID 6d9ynXJql]
10. Haun JN, Lind JD, Shimada SL, Martin TL, Gosline RM, Antinori N, et al. Evaluating user experiences of the secure messaging tool on the Veterans Affairs' patient portal system. J Med Internet Res 2014;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
11. Haun J, Noland DV, Graham-Pole J, Rienzo B, Donaldson P. Testing a health literacy screening tool: Implications for utilization of a BRIEF health literacy indicator. Fed Pract 2009;26(12):24-31.

12. Haun J, Luther S, Dodd V, Donaldson P. Measurement variation across health literacy assessments: Implications for assessment selection in research and practice. J Health Commun 2012;17 Suppl 3:141-159. [doi: 10.1080/10810730.2012.712615] [Medline: 23030567]
13. Bunz U. The Computer-Email-Web (CEW) Fluency Scale: Development and validation. Int J Hum Comput Interact 2004;17(4):479-506.
14. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
15. Dudas RA, Crocetti M. Pediatric caregiver attitudes toward email communication: Survey in an urban primary care setting. J Med Internet Res 2013;15(10):e228 [FREE Full text] [doi: 10.2196/jmir.2738] [Medline: 24152542]
16. Lee JL, Choudhry NK, Wu AW, Matlin OS, Brennan TA, Shrank WH. Patient use of email, Facebook, and physician websites to communicate with physicians: A national online survey of retail pharmacy users. J Gen Intern Med 2015 Jun 24:e1 (forthcoming). [doi: 10.1007/s11606-015-3374-7] [Medline: 26105675]
17. Cutrona SL, Roblin DW, Wagner JL, Gaglio B, Williams AE, Torres SR, et al. Adult willingness to use email and social media for peer-to-peer cancer screening communication: Quantitative interview study. JMIR Res Protoc 2013;2(2):e52 [FREE Full text] [doi: 10.2196/resprot.2886] [Medline: 24287495]
18. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: A scoping review. J Am Med Inform Assoc 2011;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
19. Kim K, Nahm E. Benefits of and barriers to the use of personal health records (PHR) for health management among adults. Online J Nurs Inform 2012;16(3):1-9 [FREE Full text]
20. Carrión SI, Fernández-Alemán JL, Toval A. Are personal health records safe? A review of free web-accessible personal health record privacy policies. J Med Internet Res 2012;14(4):e114 [FREE Full text] [doi: 10.2196/jmir.1904] [Medline: 22917868]

Abbreviations

CEW: Computer-Email-Web Fluency Scale
eHEALS: eHealth Literacy Scale
REALM: Rapid Estimate of Adult Literacy in Medicine
STOFHLA: Test of Functional Health Literacy in Adults-Short Form
VA: Department of Veterans Affairs
VAMC: VA Medical Centers

Edited by G Eysenbach; submitted 23.09.15; peer-reviewed by C Turvey, K Nazi; comments to author 17.10.15; revised version received 30.10.15; accepted 01.11.15; published 21.12.15.

Please cite as:
Haun JN, Patel NR, Lind JD, Antinori N
Large-Scale Survey Findings Inform Patients' Experiences in Using Secure Messaging to Engage in Patient-Provider Communication and Self-Care Management: A Quantitative Assessment
J Med Internet Res 2015;17(12):e282
URL: http://www.jmir.org/2015/12/e282/
doi:10.2196/jmir.5152
PMID:26690761

Original Paper

From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities

Bradley Carron-Arthur[1], BPsych (Hons); Kathina Ali[1], MSc (Psych); John Alastair Cunningham[1,2], PhD; Kathleen Margaret Griffiths[1], PhD

[1]National Institute for Mental Health Research, Research School of Population Health, The Australian National University, Acton, Australia
[2]Centre for Addiction and Mental Health, Toronto, ON, Canada

Corresponding Author:
Bradley Carron-Arthur, BPsych (Hons)
National Institute for Mental Health Research
Research School of Population Health
The Australian National University
63 Eggleston Road
Acton, 2601
Australia
Phone: 61 61256825
Fax: 61 61250733
Email: Bradley.Carron-Arthur@anu.edu.au

Abstract

Background: Understanding how people participate in and contribute to online health communities (OHCs) is useful knowledge in multiple domains. It is helpful for community managers in developing strategies for building community, for organizations in disseminating information about health interventions, and for researchers in understanding the social dynamics of peer support.

Objective: We sought to determine if any patterns were apparent in the nature of user participation across online health communities.

Methods: The current study involved a systematic review of all studies that have investigated the nature of participation in an online health community and have provided a quantifiable method for categorizing a person based on their participation style. A systematic search yielded 20 papers.

Results: Participatory styles were classified as either multidimensional (based on multiple metrics) or unidimensional (based on one metric). With respect to the multidimensional category, a total of 41 different participation styles were identified ranging from Influential Users who were leaders on the board to Topic-Focused Responders who focused on a specific topic and tended to respond to rather than initiate posts. However, there was little overlap in participation styles identified both across OHCs for different health conditions and within OHCs for specific health conditions. Five of the 41 styles emerged in more than one study (Hubs, Authorities, Facilitators, Prime Givers, and Discussants), but the remainder were reported in only one study. The focus of the unidimensional studies was on level of engagement and particularly on high-engaged users. Eight different metrics were used to evaluate level of engagement with the greatest focus on frequency of posts.

Conclusions: With the exception of high-engaged users based on high post frequency, the current review found little evidence for consistent participatory styles across different health communities. However, this area of research is in its infancy, with most of the studies included in the review being published in the last 2 years. Nevertheless, the review delivers a nomenclature for OHC participation styles and metrics and discusses important methodological issues that will provide a basis for future comparative research in the area. Further studies are required to systematically investigate a range of participatory styles, to investigate their association with different types of online health communities and to determine the contribution of different participatory styles within and across online health communities.

(J Med Internet Res 2015;17(12):e271) doi:10.2196/jmir.4705

KEYWORDS

online health community; participation style; social network; participation inequality; systematic review

Introduction

Participation rates of people in online communities are known to be highly variable with some people contributing much more than others. Across all types of online communities, the variability in degree of user participation consistently follows a pattern [1]. In particular, this pattern in participation is described by a power law. This power law means, for example, that the top 1% of participants contribute as much as 75% of the posts in an online health community (OHC) [2,3]. This pattern is indicative of a coherent community [2], and these highly engaged individuals are repeatedly observed in well-established OHCs [4]. These individuals are of interest. Their high participation rates and predictable presence suggest that they may be of particular value to the OHC.

Although post frequency may constitute a simple indicator of engagement, from post frequency alone it is not possible to ascertain exactly what ways a person contributes. Post frequency does not indicate whether a person starts new discussions, welcomes newcomers, is available at critical times in the day when people are most likely to need support, or is knowledgeable about certain topics. In order to ascertain whether people contribute these different kinds of value, it is necessary to measure their participation based on various other metrics.

There may be value for those who are involved in the development of an OHC to identify users who contribute particular types of value to the OHC. This points to the need for multiple metrics to define user contributions. For example, in a qualitative paper on building and sustaining OHCs, Young described how certain core members were vital to the development and sustainability of an OHC [5]. As the community manager from the inception of this OHC, Young was able to provide an account of the different ways that these users had contributed to the development of the OHC including facilitating discussion and fostering a supportive culture. Young also suggested ways that OHC managers might harness the contributions of these individuals to help build the community by, for example, highlighting their best posts or inviting them to contribute to a community resource such as a newsletter.

For a variety of reasons, including time constraints and size of the community, not all community managers are able to have a strong qualitative understanding of the roles of particular individuals in their OHC. However, community managers would potentially benefit from a simple operationalization of user participation in terms of metrics that are automatically collected in the log data of the OHC software. This would help them to identify the core members and various other users who contribute in different ways so that they may apply the community building techniques recommended by Young [5].

OHCs also provide an opportune setting for interventions that encourage certain positive health behaviors [6]. Knowing who the most influential people are in an OHC, or how to reach most of the community via the smallest subset of people, might inform dissemination activities such as promoting new evidence-based treatments or recommending correct use of certain medications. Finally, there is scientific value in investigating the ways in which different people participate in OHCs across multiple contexts. There may be patterns in the way in which people participate that can be found across multiple different OHCs. These patterns may help us learn more about the social dynamics of OHCs and the way that people seek help and provide it to others.

User profiling by categorizing participation styles is conducted in studies of online communities more broadly. There are some roles such as “newbies” and “celebrities” that may be found in any online community, but most others are likely to be specific to the type of community [7]. For example, “technical editors” and “substantive experts” are found in Wikipedia [8], but these may not be relevant to or found in OHCs. We expect that OHCs will have high-profile users who are akin to “celebrities,” but the nomenclature and the metrics used to define these users may be tailored to the supportive context and health discussion focus of the community. There may be further similarities and differences between participation styles in communities of different health types.

This study seeks to advance this area by conducting a systematic review of all studies that provide replicable, quantifiable criteria for categorizing the nature of participation in an OHC. We aimed to document all participation styles that had been identified to date and the OHCs from which they came. Our objective was to determine if any patterns were apparent in the nature of user participation across OHCs for different health conditions or within each.

Methods

A systematic review was conducted to identify articles that investigated participation styles in an online health community. For the current purposes, an online health community was defined as any Internet-based platform designed to enable people to communicate about health issues. A participation style was defined as any type of engagement with an OHC that can be measured quantitatively. This does not include simply the presence or absence of participation (ie, posters and lurkers), as this has been well documented elsewhere [9], but rather is aimed at understanding the nature of participation for those who are actively engaged in the community.

Search Strategy

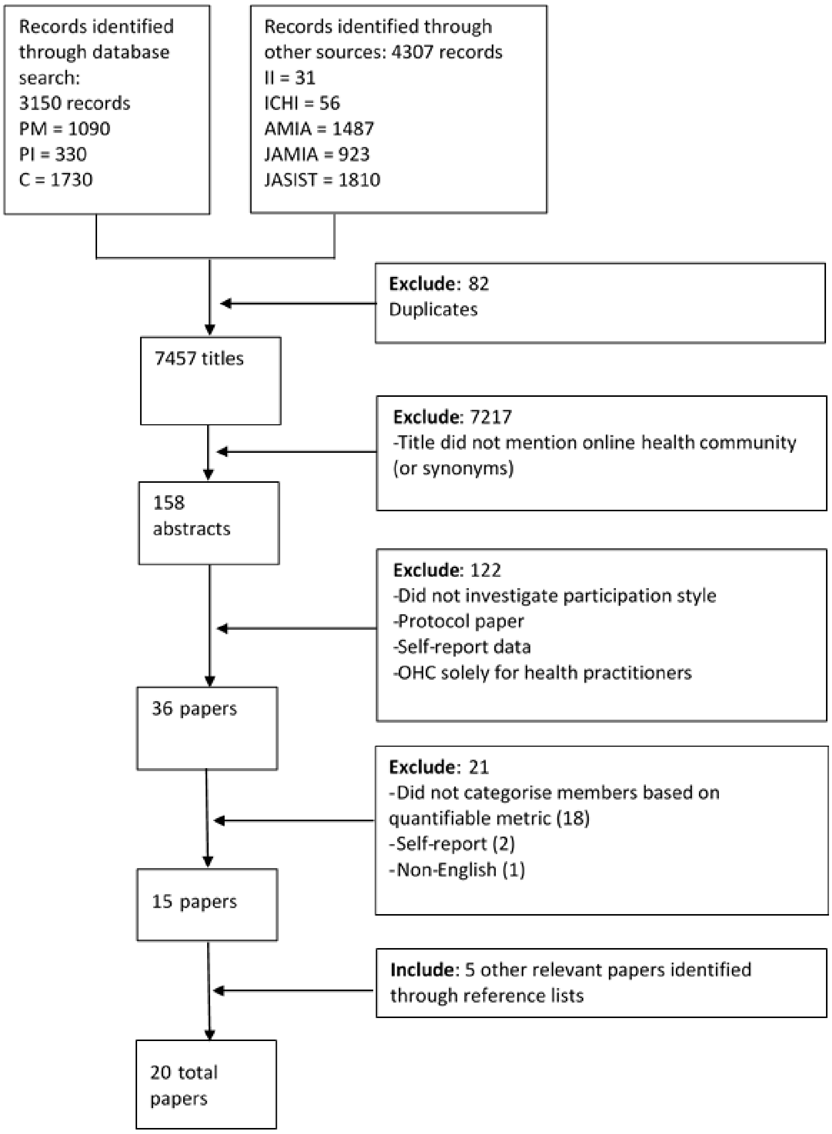
Three databases (PubMed, PsycINFO, and Cochrane) were searched for all articles prior to December 2014. Adapted search terms from Eysenbach et al [10] and Griffiths et al [11] were used to identify the concept of OHC (see Multimedia Appendix 1). These search terms were combined with the following terms to identify the participation style concept: (participatory pattern*) OR (posting pattern*) OR (posting behavior pattern*) OR (use pattern*) OR (communication pattern*) OR (usage pattern*) OR (system use*) OR (traffic) OR (participative stance*) OR (participant contribution*) OR (posting habits*) OR (participation rate*) OR (posting rate*) OR (user engagement) OR (level* of engagement*) OR (pattern* of engagement*) OR (type* of engagement) OR (share information) OR (community structure) OR (social dynamics).

In addition, papers from relevant journals and conference proceedings in the computer and information science field published since 2005 (including the *American Medical Informatics Association Annual Symposium*, *Journal of the American Medical Informatics Association*, *Journal of the Association for Information Science and Technology*, and *International Conference on Healthcare Informatics*) and a new journal that was not yet indexed at the time of the search (*Internet Interventions*) were screened for relevant articles.

Article Selection

A total of 7457 articles were screened. Of these, 3150 were retrieved from the database search and 4307 were from the additional journals and conference proceedings. A total of 82 duplicate articles were identified and removed. Relevant articles were selected through a multistage process (Figure 1). Initially, titles were screened by 2 raters (BC and KA). Any article that mentioned an online community or synonym thereof in the title (or online health community in the case of the *Journal of the Association for Information Science and Technology*) was included. This reduced the number of articles to 158. The abstracts of these articles were subsequently screened by the 2 raters. Any article that investigated ways that people participate in an online health community was included. Articles based on self-report measures of OHC use and research protocols were excluded. The full articles for the 36 remaining abstracts were retrieved and read by both raters. Any disagreements between the raters were resolved by discussion.

Figure 1. Study identification flow diagram: PubMed (PM), PsychINFO (PI), Cochrane (C), Internet Interventions (II), International Conference on Healthcare Informatics (ICHI), American Medical Informatics Association Annual Symposium (AMIA), Journal of the American Medical Informatics Association (JAMIA), Journal of the Association for Information Science and Technology (JASIST).



Inclusion Criteria

The final set of articles included any study that (1) quantitatively investigated ways that people participate in an online health community, and (2) categorized users based on any quantifiable metric that can be used to show they have engaged with the community.

Exclusion Criteria

Studies that converted written content to quantitative data by a means that was computerized (eg, machine learning algorithm) were included, but studies that relied on human interpretation of written content to create quantitative data were not. This ensured that the methods identified could be accurately replicated and would be scalable to large OHCs. For similar reasons, studies that used self-report data from surveys were not included. This meant that only studies reporting data that had been automatically logged by the OHC software or that had been extracted by programs that crawl publicly available data were included in this systematic review. Protocol papers, articles not written in English, and papers on OHCs solely for health practitioners were not included.

After applying these criteria, a set of 15 papers were included. The reference lists of included papers and those that cited them (as per Google Scholar) were hand searched. This yielded an additional 5 papers, resulting in a final set of 20 included papers.

Coding

The included papers were coded by 1 rater (BC). Each participation style identified by a paper was listed. Three attributes of each participation style were coded: (1) the name used by the authors to describe the participation style, for example, “superuser,” (2) the metrics used to quantitatively describe their style of participation, for example, frequency of posts, and (3) the inclusion criteria used to determine who was categorized as having that participation style, for example, the top 1% of users whose frequency of posts was greatest were deemed to be superusers.

Results

Across the final set of 20 papers, users were categorized into participation styles a total of 74 times, of which 28 were duplicates. These duplicates included participation styles that had been assigned different names by different studies but used the same metrics and same inclusion criteria (or very similar) to define them. By merging all these redundant categorizations into the same participation style, we determined that 44 participation styles had been identified in OHCs to date.

Table 1 [2,3,12-29] shows a summary of information about the OHCs where the participation styles were identified. Some studies investigated more than one OHC. In total, there were 26 different OHCs. These were used for a variety of different health topics including smoking cessation (n=7), cancer (n=6), mental health issues (n=6), diabetes (n=5), multiple sclerosis (n=1), and social innovation in health care (n=1). These OHCs were based in different countries including the United States (n=8), Canada (n=2), Australia (n=1), Germany (n=1), New Zealand (n=1), Norway (n=1), Taiwan (n=1), and the United Kingdom (n=1). The country of origin for 10 OHCs was not reported. The sample of people drawn from each OHC ranged in size from 77 to 49,552 people. Most included between 1000 and 10,000 people; however, one group of 5 OHCs included more than 140,000 people between them. All of the studies were published in 2007 or later, with 12 of the 20 published since 2013.

Table 2 [2,3,12-29] shows a summary of these types of participation. Within Table 2, we have grouped participation styles first into two categories: those based on multiple metrics (multidimensional) and those based on one metric (unidimensional). Each of these is also then divided into up to 3 categories according to the predominant type of metric used to define the participation style: activity-based, network-based, and content-based metrics. Table 3 [30,31] contains a list of the metrics and a description of what they measure.

There were 41 participation styles in the multidimensional category (13 activity based, 11 network based, and 17 content based). In all instances where a unidimensional participation style was identified, the studies divided the users into no more than 3 groups that we have summarized as high, medium, and low engagement. There were 8 different metrics used in the high engagement category (5 activity based, 3 network based), 3 in the medium category (2 activity based, 1 network based), and 4 in the low category (3 activity based, 1 network based).

The results of each subcategory of participation style (content based, network based, and activity based) are described in turn for the 41 multidimensional participation styles. Following this, the results of the unidimensional participation styles are described together for each of the 3 participation styles identified.

Table 1. Summary of online health community characteristics.

Online health community name	Year of study	Health condition	Country	Sample size, n
SOL-Cancer Forum	2007 [27]	Cancer	Not reported	84
Cancer Survivors Network	2014 [22,23]	Cancer	United States	27,173
Cancer Compass	2011 [28]	Cancer	United States	7991
WebChoice	2013 [13]	Cancer (breast and prostate)	Norway	103
Breastcancer.org	2014 [29]	Cancer (breast)	United States	49,552
Cancer Compass	2010 [17]	Cancer (melanoma)	United States	851
Five unnamed forums in English and Spanish	2013 [14]	Diabetes	Not reported	>140,000
BlueBoard	2014 [2]	Mental health	Australia	2932
DepressionCenter	2014 [3]	Mental health (depression)	Not reported	5151
PanicCenter	2014 [3]	Mental health (panic disorder)	Not reported	11,372
AlcoholHelpCenter	2014 [3]	Mental health (problem drinking)	Not reported	2597
PTT.CC—Psychosis Support Group	2009 [26]	Mental health (psychosis)	Taiwan	438
SharpTalk	2011 [19]	Mental health (self-harm)	United Kingdom	77
Deutsche Multiple Sklerose Gesellschaft	2014 [20]	Multiple sclerosis	Germany	1169
The Canadian Cancer Society's Smokers' Helpline Online	2012 [21]	Smoking	Canada	1670
QuitBlogs	2014 [18]	Smoking	New Zealand	3448
Alt.Support.Stop-Smoking	2014 [25]	Smoking	Not reported	8236
QuitPlan	2008 [12]	Smoking	United States	233
QuitNet	2010 [15]	Smoking	United States	7569
	2013 [16]			Not reported
StopSmokingCenter	2012 [21]	Smoking	United States	1627
	2014 [3]			44,870
#HCSMCA	2013 [24]	Social innovation in health care	Canada	486

Table 2. Summary of participation styles including name, metrics, and inclusion criteria.

Name	Metrics	Inclusion criteria
Multidimensional		
Content based		
Influential user [23]	69 activity, network, and content features including influential responding replies	Machine learning classifier (relying initially on expert judgement to identify exemplars)
Leader [22]	68 activity, network and content features	Machine learning classifier (relying initially on expert judgement to identify exemplars)
Opinion leader [16]	Word vectors, degree	Latent semantic analysis and high degree
Information providers [29]	Social support type	High information support
Community builders [29]	Social support type	High companionship support
Emotional support providers [29]	Social support type	High emotional support
Information seekers [29]	Social support type	High information support seeking
Emotional support seekers [29]	Social support type	High emotional support seeking
Information enthusiasts [29]	Social support type	High information support seeking, high information support
All-around contributors [29]	Social support type	No particular metric stands out
Balanced source user [20]	Source of information	Cited information from a range of sources
Social media fan [20]	Source of information	High social media
Organization follower [20]	Source of information	High organizations
Homepage promoter [20]	Source of information	High static informational websites
Seeker of health care [20]	Source of information	High health practitioners
User of uncommon sources [20]	Source of information	High uncommon sources
Sophisticated contributor [20]	Word count, source of information	High word count, high academic references
Network based		
Key player [15]	Degree (nonredundant)	Key Player 1.4 software
Hub [14,17,28]	Out-degree, in-degree	Hyperlink-induced topic search algorithm
Authority [14,17,28]	Out-degree, in-degree	Hyperlink-induced topic search algorithm
Facilitator [17,28]	Out-degree, in-degree	Hyperlink-induced topic search algorithm
Trusted user [14]	Out-degree, in-degree	PageRank algorithm
Help-seeker [14]	Out-degree, in-degree	Low in-degree, high out-degree (within the scope of the edge between 2 users)
Star [27]	Out-degree, in-degree	Top ranked individual (outlier)
Prime givers [14,27]	Out-degree, in-degree	Very high out-degree, high in-degree
Serious members [27]	Out-degree, in-degree	Moderate out-degree, moderate in-degree
Moderate users [27]	Out-degree, in-degree	Low out-degree, low in-degree
Takers [27]	Out-degree, in-degree	No out-degree, low in-degree
Activity based		
Caretaker [19]	Time logged in, episodes, reading, posting, thread initiation	High time logged in, low episodes, high reading, low posting, low thread initiation
Here for you [19]	Thread initiation, posting, forum	Low thread initiation, high posting in support forum
Butterfly [19]	Time logged in, episodes, posting, forum	High time logged in, high episodes, high posting in support forum

Name	Metrics	Inclusion criteria
Crisis-oriented individual [19]	Posting, forum	High posting in support forum
Discussant [19,20]	Thread initiation, posting, forum	High thread initiation, high posting in discussion forum
Average user [20]	Thread initiation, posting, forum, topic, days active, word count, source of information	No particular metric stands out
Highly active relational poster [20]	Posts per day, thread participation, thread initiation	High posts per day, high thread participation, low thread initiation
Topic-focused responder [20]	Thread initiation, posting, topic, days active	Low thread initiation, low posts per day, high fraction of topic-related posts, low days active
Topic-spammer [20]	Posting, days active, word count, topic, source of information	Low days active, high posting, low word count, high fraction of topic-related posts, low references
Long-term high-activity users [25]	Days active, posting	High days active, high posting
Short-term high-activity users [25]	Days active, posting	Low days active, high posting
Short-term low-activity users [25]	Days active, posting	Low days active, low posting
Long-term low-activity users [25]	Days active, posting	High days active, low posting
Unidimensional		
Activity based		
High-engaged user	Posting	>2 posts [12]; top 1% of users [2,3]; top 10 users [24]; >180 posts [18]; top 100 users [21]
	Reading	>5 posts [12]
	Time logged in	Top 33.3% of users [13]
	Thread initiation	Top 100 users [21]
	Thread participation	Top 100 users [21]
	Network based	
	Friendship	Mutual friend nomination between 2 users and >4 interactions between them [15]
	In-degree	Top 10 users [24]; high in-degree [26]
	Out-degree	Top 10 users [24]; high out-degree [26]
Moderate-engaged user	Posting	2-10 percentile (9%) of users [2,3]
	Time logged in	Middle 33.3% of users [13]
	Network based	
	Friendship	Friend nomination of another user and >0 interactions with them [15]
Low-engaged user	Posting	1-2 post [12,18]; bottom 90% of users [2,3]
	Reading	1-5 posts [12]
	Time logged in	Bottom 33.3% of registered users [13]
	Network based	
	Friendship	Any interactions with another user [15]

Table 3. A description of the mtrics used to classify participation styles.

Metric	Description
Activity-based metrics: measure the individual actions taken by users in an OHC	
Posting	Number of posts a person has made in the OHC
Time logged in	Amount of time a person has spent accessing the OHC
Reading	Number of posts that a person has read
Thread initiation	Number of times a person has created a thread
Episodes	Number of times a person has accessed the OHC
Days active	Number of days between a person's first and last post
Forum	Number of posts a person has made in a particular subforum of the OHC, eg, support or discussion
Thread participation	Number of different threads a person has posted in
Network-based metrics: measure the relationship and interactions between users	
Degree (in/out)	The number of people a person has communicated with. Where it is possible to tell who the source of the communication was and to whom it was directed, the number of people a person has made outgoing communication with is called the "out-degree" and the number of people that a person has received communication from is called the "in-degree." When it is not possible to tell the direction, the communication is counted for both people as a measure of degree. Degree is considered to be a measure of a user's centrality in a network [30,31].
Friendship	The extent to which a person is connected with at least one other person in the OHC as defined by 3 thresholds: Low—any interactions with another user; Moderate—friend nomination of another user and >0 interactions with them; and High—mutual friend nomination between 2 users and >4 interactions between them.
Content-based metrics: measure the nature of the content within posts	
Word vectors	A representation of the proportion of words in a message that fit a certain topic.
Influential Responding Replies	Number of posts a person has made that have influenced the sentiment of the thread initiator
Social support type	Number of posts a person has made that either provide or seek information support, emotional support, or companionship
Topic	Number of posts a person has made which included subject matter on a specific topic
Source of information	Number of citations a person has made from a particular source
Word count	Number of words in a post

Multidimensional

Content-Based

Leaders and Influential Users

Zhao et al [23] created a machine learning classifier with 69 metrics that was used to identify influential users in an OHC. These users were regarded as leaders who could influence the emotional sentiment of other users. This study built on previous research by Zhao et al [22], which used 68 metrics such as number of posts, in-degree, and days active in a classifier to first identify leaders in the OHC. Zhao et al [23] then created a metric called "influential responding replies (IRRs)." This was the number of times a person was able to affect the sentiment of another person when responding to their initial post. It was found that this metric alone outperformed the classifier with 68 metrics, and together they created the best performing classifier. In order to train this IRR-enhanced classifier, it was necessary to have a list of users who were deemed to be influential users by moderators of the OHC. There were 41 users in this list. In total, the moderators identified 126 influential users. A list of the top 50, 100, and 150 influential users identified by the classifier was made up with 90%, 77.7%, and 68.7% users from the moderator list of 126 respectively. The highest percentage possible in the 150 influential user condition was 84.0% (126/150).

Opinion Leaders

Myeni et al [16] used latent semantic analysis to identify users who were involved in discussion about particular concepts such as personal experiences, advice, or adherence to interventions. Users whose mean word vector scores for a concept were one standard deviation above the sample mean were grouped together in a social network. Within each theme-based social network, Opinion Leaders were identified as people who had the highest degree. These people were considered to be influential in their specific domain and may be particularly useful to identify when delivering relevant targeted interventions. Subsequent research has shown that exposure to users who were abstinent from smoking in the theme-based networks of "social support" and "traditions" were more likely to be abstinent themselves [32].

Information Providers, Emotional Support Providers, Information Support Seekers, Emotional Support Seekers,

Community Builders, Information Enthusiasts, and All-Around Contributors

Wang et al [29] created a machine learning classifier to determine which posts in a cancer OHC with more than 2.8 million posts contained each of the following types of content: providing informational support, providing emotional support, seeking informational support, and seeking emotional support and companionship. The authors then used a k-means clustering algorithm [33] to categorize all users based on the proportion of posts they made with each type of content. This produced 7 types of users. Five were typified by writing a high proportion of posts that predominantly contained one type of each of the above 5 content types. The remaining two, information enthusiasts and all-around contributors, were typified by having equally high proportions of posts seeking and providing informational support, and having equal amounts of all types of content, respectively. The all-around contributor was the most common type of user of all 7 (making up 32% of all users). Community builders were among the least common (8%) but were responsible for writing the most posts on average along with all-around contributors. Those who primarily engaged in informational and emotional support types posted less and did not remain in the forum as long as community builders and all-around contributors.

Balanced Source User, Social Media Fan, Organization Follower, Homepage Promoter, Seeker of Health Care, and User of Uncommon Sources

Sudau et al [20] observed that people tend to favor different sources of information to support the points that they make in posts. A number of participation styles represent this bias. In order to determine these participation styles, Sudau et al used a k-means clustering algorithm [33] to form 6 groups of similar users based on the frequency of different hyperlinks they used from 8 domain classes. The groups were labeled according to what Sudau et al thought best described their referencing tendencies.

Sophisticated Contributor

A sophisticated contributor is a user whose posts are longer than those of the average user participation style and contain more references. In contrast to the activities of most users, these references are more often to scientific publications than to social media sources. Sudau et al [20] identified this participation style in 4 of 171 users. Sophisticated contributor posts were three times as long and contained five times as many references as posts by Average Users.

Network Based

Key Players

Cobb et al [15] sought to identify a set of users who were maximally connected to other users throughout the social network of the OHC. A set of key players is a small group of a specified number of users who are connected with as many other people in the network as possible, for example, through private message, posting, or friendship. Cobb et al used Key Player 1.4 software [34] to determine the reach of a set of 50 key players. These 50 key players were connected to 64% of other users in the network. Note that these are not necessarily the 50 most connected individuals in the OHC; that is, they are not the top 50 users ranked by degree. Rather, the algorithm considers redundancy. If introducing a new key player to the set does not increase the set's overall reach, that player is not added. The optimum key player set of 50 users may not necessarily contain all the users in the 49 set nor will either necessarily contain the user who, as an individual, is the most connected person in the network. The intention of the algorithm is to enable maximum access to the whole network through minimal nodes. This, for example, enables maximum efficiency in dissemination of information.

Hubs, Authorities, and Facilitators

Hubs and authorities are concepts borrowed from the computer science literature on the Web. Hubs and authorities are identified using the hyperlink-induced topic search (HITS) algorithm [35]. In this algorithm, every website receives both a hub and an authority score. High-scoring authorities are websites that are linked to high-scoring hubs. High-scoring hubs are websites that link to high-scoring authorities. Websites with high authority scores tend to those that provide good information on a specific topic. Hubs direct people to these various authorities. The algorithm can be applied to any network consisting of nodes and links between them by analyzing the pattern of out-degree and in-degree across the network. Accordingly, both Chomutare et al [14] and Durant et al [17,28] have used the HITS algorithm to identify people in OHCs as authorities and hubs. The 3 papers have posited that those identified as hubs are people who disseminate information by promoting discussion. They have a relatively high out-degree in the network compared with their in-degree. They are important for sustaining the activity levels of the community. Authorities are people whose opinion is highly respected in the community. They have a relatively high in-degree. A third participation style—a facilitator—was also proposed by Durant et al [17]. A facilitator is a person who is ranked similarly highly as a hub and as an authority. They are considered to be more effective for sustaining communication in the OHC than those who are hubs or authorities alone. Durant et al [28] sought to track the presence of facilitators over time by segmenting and analyzing the network each year over an 8-year period and found that the top 5% were rarely the same individuals in consecutive years.

Trusted Users

Similar to the HITS algorithm, the PageRank algorithm [36] is another method originating in the computer science literature on the Web. Rather than identifying 2 types of users, the PageRank algorithm identifies one type. The score given to each node in the network by the PageRank algorithm is the probability of arriving at that node given a random walk around the network via the links between them. This means that nodes that are linked to more often have higher probabilities of being landed on, and nodes that are linked to more often by other high-scoring nodes have even higher scores. If it is assumed that a directional link between 2 nodes is a vote of support to the other, this algorithm identifies trusted users. This algorithm was the basis for Google search. Chomutare et al [14] have applied it to an OHC and have made the same assumption. They found that 6 out of 10 of the highest ranked users by in-degree were also in the top 10 identified by the PageRank algorithm.

Help-Seekers

In a relationship between 2 people where one communicates with the other much more often, the person who instigates more communication (higher out-degree than in-degree) is labeled a Help-Seeker. Chomutare et al [14] suggested that this pattern of metrics might reflect a person who is struggling with their health issue. However, the authors noted that the nature of the help-seeking is not exactly clear as the user may either be strongly motivated to engage in self-care or they may be a particularly needy user, and neither can be concluded without content analysis. The authors originally suggested the label “needy user” for this participation style, but we have renamed it “help-seeker” given the ambiguity and lack of clarity around the concept of needy in this context.

Star, Prime Givers, Serious Members, Moderate Users, and Takers

The earliest recorded participation styles were identified by Bambina [27] who compared the in-degree and out-degree of users and grouped them around a pattern in the results that was related to engagement. Bambina first noted one outlier: a person who had both the highest in-degree and out-degree. Bambina referred to this person as the “star.” This person provided the most support to others including notably many new individuals with whom many others did not communicate. Bambina noted that the next most engaged people by both in-degree and out-degree all tended to provide more support than they received, that is, have higher out-degree than in-degree. These were named “prime givers” (n=6). Chomutare et al [14] observed the same pattern in a social network analysis that they conducted, but they did not report whether it was associated with providing support. Bambina also noted 2 groups who had relatively similar in-degree and out-degree within each group. These were the designated “serious members” (10) and “moderate users” (n=15). Last was a group labeled the “takers” who never provided support but who initiated a conversation and received support from others (n=52).

Activity-Based

Caretakers

Jones et al [19] identified one user in a sample of 77 people as having a participation style called the “caretaker.” They identified this person, as they did for all participation styles, through visual inspection of scatterplots of various metrics. The OHC was a support group for young people who self-harm. Given the large amount of time the person spent logged in, they actively participated very little. The times they did post were largely in response to other users rather than initiating their own threads. Jones et al concluded that this person might be watching over the whole forum and looking out for others in need. This person undertook the caretaker role despite the OHC being a moderated forum.

Here for You

One user in a sample of 77 people was considered to take the “here for you” participation style by Jones et al [19]. Like the caretaker, they did not create many threads of their own. However, in contrast to the latter, they did post large amounts of comments in response to other people who needed support.

Butterfly

Another user in the Jones et al [19] sample was classified as being characterized by a butterfly participation style. This person logged in many more times than anyone else. They spent short amounts of time checking out a few pages and then logged out again. They posted mostly in the support forum (as opposed to the discussion forum or off-topic forum). Like the crisis-oriented individuals in the following section, they were considered by the moderators to be in crisis and needing support as opposed to providing it.

Crisis-Oriented Individuals

Six users of the Jones et al [19] sample were classified as crisis-oriented in their participation style. These people posted in large numbers in the support forum. It is not possible to confirm from the objective metrics alone whether such people were in crisis or providing support; however, it was confirmed by the moderators of the forum that all 6 were indeed in crisis. These users did not visit the OHC as frequently as the user with the butterfly participation style.

Discussants

A discussant is a user who is mainly focused on discussion about health-related topics as opposed to providing or receiving support. They initiate a high number of threads in the discussion section of the OHC and participate actively in them. This participation style was identified by both Jones et al [19] and Sudau et al [20].

Average Users

A user type that is not distinctly based on any metric, the average user category was identified by the application of a second k-means clustering algorithm conducted by Sudau et al [20]. This analysis was designed to form 6 groups of similar users based on 9 metrics that measured their active participation in the community. Sudau et al labeled the groups according to their distinguishing features. Average users were a group that were thought not to exhibit any distinguishing features. This group constituted 63% of the people included in the analysis.

Highly Active Relational Posters

These are the most active users of an OHC by post frequency. Sudau et al [20] noted these users maintain “small talk,” which may be good for community building. They participate in many different threads but do not initiate many themselves.

Topic-Focused Responders

A user whose activity is concentrated on a specific topic, the topic-focused responder is distinct from a discussant in that they do not post as much and do not initiate as many threads. Sudau et al [20] included only people who had made at least five posts on a certain topic in their analysis. Topic-focused responders met this criterion but they did not have many other posts. They tended to focus mainly on responding to others who had initiated the topic. Sudau et al suggested this style may be similar to the here for you participation style identified by Jones et al [19], but we have separated them because of the distinction between discussion and support.

Topic-Spammers

This is a user who is active for a very short period, that is, only a few days. In that time, they contribute a high number of posts on a specific topic in the discussion forum. However, these are not particularly sophisticated posts, rather they are short and lack references. This participation style was identified by Sudau et al [20].

Short-Term and Long-Term, High-Activity and Low-Activity Users

Stearns et al [25] noted that the bulk of users in a smoking cessation OHC are made up of short-term users (active for approximately less than 1 week), who, regardless of whether they have high or low activity, tend to be involved in the OHC for personal gain. Long-term users with low activity are noted to have smaller social circles and a stronger interest in particular topics. Stearns et al state that long-term high-activity users are most like Young's [5] "core members" who are vital to the sustainability of the OHC.

Unidimensional

All but one of the studies [24] that made unidimensional classifications did so for the purpose of determining if the type and level of engagement a person showed was predicted by demographic factors and whether high engagement predicted specific health outcomes. Some studies made statements about the nature of participation of users in the OHC. Given that the purpose of this review was to investigate the nature of participation, we focus on reporting these findings in the following sections considering first high-engaged users, followed by moderate- and low-engaged users.

High-Engaged Users

All 8 studies that made a unidimensional categorization [2,3,12,13,15,18,21,24] classified users into a participation style that we call high-engaged users. There were 8 different metrics used across these studies that all indicate a different type of high engagement. These included posting frequency, thread initiation, thread participation, level of in-degree/out-degree, reading of posts, time logged in, and friendship (see Table 3 for definitions).

Frequency of posting was the most commonly used metric used by 6 of the 8 studies [2,3,12,18,21,24]. It was used to classify users in a total of 9 OHCs, with 4 of those being for smoking cessation, 4 for mental health issues, and 1 for social innovation in health care. Users who were highly engaged according to posting frequency were regarded by all but one of the studies [12] as being valuable to the OHC because they sustained activity levels and in doing so facilitated the engagement of others. Four of the 6 studies referred to these people as either "superusers" [2,3,21] or "community leaders" [24]. This regard spanned across all the types of OHCs mentioned earlier.

Thread initiation and thread participation (together with posting frequency) were used by one study [21] to classify the top 100 ranked users, denoting them "superusers." The moderators of the OHC were asked to identify leaders within it. The authors noted that although most studies have previously identified leaders in an OHC using posting frequency alone, the moderators thought it was necessary to also include users who start many conversations and who participate in many different conversations in their definition of a "superuser."

In-degree and out-degree were employed by 2 studies to classify users as highly engaged [24,26]. The authors of one study [24] regarded users with high in-degree (top 10) as authorities on topics, similar to the hubs and authorities discussed earlier. This study was conducted on an OHC that existed within Twitter. It was noted that those people with the highest in-degree were also people who had the highest number of followers on Twitter in general. Users with high in-degree were considered to be valuable for engaging other less active users in discussion. It was noted that the 6 users were both top 10 ranked users by in-degree and out-degree. These 6 people were thought to be communicating on topics that resonated with the community and were considered to be "community leaders." In a study of a mental health OHC for psychosis, Chang et al [26] referred to users with either a high in-degree or out-degree as "stars" after Bambina's [27] single outlying user.

Other metrics employed to classify users as highly engaged included reading [12], time logged in [13], and friendship [15].

Moderate-Engaged Users

Four studies classified users as moderately engaged based on 3 different metrics. Two were based on posting frequency [2,3], and one each on time logged in [13] and friendship [15].

Low-Engaged Users

Six studies classified users as low engaged based on 4 different metrics. Four were based on posting frequency [2,3,12,18], and one each on reading [12], time logged in [13], and friendship [15].

Discussion

Principal Findings

This systematic review synthesized findings from studies that investigated the nature of participation in an OHC by categorizing users based on metrics of participation. The aim of this review was to identify the different ways in which users participate and contribute to OHCs, although we acknowledge that the resultant list of participation styles may not provide a comprehensive account of all possible styles. Our objective was to determine whether any patterns were apparent in the types of participation styles that were identified across and within different health conditions. With the exception of an overlap in engagement measured by posting frequency (which has been discussed elsewhere [3]), there was little overlap in participation styles identified across OHCs for different health conditions or within OHCs for specific health conditions. Consequently, it is not possible for this study to address this objective. This area of research is in its infancy, with most of the studies included in this review being published in the last 2 years. Despite this shortcoming, the current review delivers a nomenclature for OHC participation styles and metrics that will provide a basis for future comparative research in the area. To inform future research, we discuss in the following section some methodological considerations for studies seeking to replicate or expand on the methods identified by this review.

Methodological Considerations

Posting Frequency

It was common for studies to use posting frequency as the sole means of classifying highly engaged users in an OHC. It was also common among these studies for researchers to regard these users as being particularly valuable to the OHC. However, it is not possible to know from post frequency alone in what way a person is contributing to an OHC. They might be contributing trivial or critical messages or their post might in other ways fail to support others. The rationale for the inference that high engagement is synonymous with high value may relate to another commonality across papers. The authors in question were also community managers of the OHCs that they were studying; therefore, they may have based their conclusions on reading content posted by these users. However, content analysis research is required to investigate whether posting frequency is a valid means of identifying generically valuable users.

Machine Learning

Zhao et al [22,23] used a complex method of identifying the participation styles of leaders and influential users that may be subject to issues with generalizability. Ideally, the classifier would be transferable across OHCs. However, there is currently no evidence to support such transferability. Indeed by using 69 metrics in their machine learning classifier, they may have created a model that is overfitted to the data of the OHC from which it came and it may not work well at identifying leaders or influential users in other OHCs, even of the same health condition. Furthermore, an essential prerequisite for the development of the method was identifying a priori, using subjective judgments, a sample of leaders and influential users for use in the learning classifier trial. Thus, if Zhao's classifier is not generalizable, research to identify a new model requires expertise, or access to expertise, in identifying leaders and influential users through qualitative methods in addition to advanced understanding of machine learning methods. Despite these challenges, research in this area offers promise, particularly as influential users most closely resemble those vital users whom Young [5] described as core members. For those who are not inclined to build their own classifier, it is noteworthy that one particularly useful and generalizable aspect of the method for determining influential users was the discovery of the metric influential responding replies, which is defined as the number of posts a person has made that have influenced the sentiment of the thread initiator. Zhao reported that this metric was a better predictor of influential user status than the other 68 metrics combined. IRRs are determined by analyzing the degree of positive and/or negative sentiment expressed in the text. There are many existing programs that can conduct this kind of sentiment analysis, such as Linguistic Inquiry and Word Count [37]. However, note that it is important to test the validity of these programs in any new dataset by comparing human and computer ratings. As Zhao points out, the word "positive" in the context of a cancer diagnosis can be a negative concept. Applying a standard sentiment analysis program in this context would yield invalid results.

Wang et al [29] also used a machine learning classifier; however, their method may be more reliably replicated without expert knowledge. The classifier was designed to detect the presence or absence of certain types of social support in posts. They used 5 human coders to classify a sample of posts that could be used for training the classifier. These people were not domain experts. Similar research has involved contracting online Amazon Mechanical Turk workers to code the presence of social support in posts for the same purpose [38]. These people also did not have prior experience in this area.

Centrality Algorithms

Similar to IRR, some participation styles described users who were useful in a particular way that would be potentially identifiable in any OHC, or for that matter, any social network. These were based on algorithms that used measures of centrality such as in-degree and out-degree. This includes authorities, hubs, facilitators, and trusted users. While these categories are quite useful, it should be noted that these algorithms are calculated in such a way that they introduce bias based on time elapsed such that users who participate earlier in the OHC receive higher scores [39]. There are methods to adjust for this [40].

K-means Clustering and Multivariate Outliers

Other more specific participation styles described users who have particular characteristics and may be found only in a subset of OHCs. This included, for example, the caretaker or the topic-spammer. The techniques used to identify these participation styles, k-means clustering algorithms and multivariate outliers, may not necessarily identify the same participation styles in other OHCs. However, they may be useful for identifying other particular or unique ways of participating in OHCs.

Limitations and Future Research

The scope of this study is quite broad. We included all studies that categorized a type of participation in an OHC despite the possibility that the culture and nature of participation in populations with different health conditions and with or without moderators could differ markedly. There was little overlap in the use of categorizations to define particular participation styles either in OHCs broadly or within specific health conditions. Thus, it is not possible to draw many specific conclusions at this early stage. A possible limitation and reason for this is that we may not have included all relevant studies, as our search terms may not have encompassed all the different terms used to describe participation styles at this early stage of research. Nevertheless, by synthesizing the findings of the included studies, this review provides a basis for future research to investigate the validity of styles identified to date by attempting to replicate findings for specific OHCs and exploring their validity across different OHCs. Future research should also investigate new participation styles not documented in this review.

Conclusion

Our systematic review identified a range of participation styles. Some of them may be generalizable to other OHCs. Others were more specific to particular OHCs but were identified by methods that could be used elsewhere. The findings of this review are intended to support the work of community managers in building community, organizations seeking to design targeted interventions and disseminate information through certain types of people in OHCs, and researchers seeking to understand the nature of peer support. We anticipate that this review will be useful for these groups in conducting investigations to determine the presence of participation styles that may be relevant to their work. However, it is too early to draw any conclusions about which OHCs would be most likely to contain users who have specific participation styles.

Acknowledgments

BC is supported by an Australian Postgraduate Award. KA is supported by a Young and Well CRC PhD Scholarship. KG is supported by the Australian National Health and Medical Research Council (NHMRC) Research Fellowship .

Conflicts of Interest

None declared.

Multimedia Appendix 1

OHC concept search terms.

[PDF File (Adobe PDF File), 17KB - jmir_v17i12e271_app1.pdf]

References

1. Nielsen Norman Group. Participation inequality: lurkers vs contributors in internet communities URL: http://www.nngroup.com/articles/participation-inequality/ [accessed 2014-06-06] [WebCite Cache ID 6Q7EwEncA]
2. Carron-Arthur B, Cunningham JA, Griffiths KM. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. Internet Interventions 2014 Oct;1(4):165-168. [doi: 10.1016/j.invent.2014.09.0033]
3. van Mierlo T. The 1% rule in four digital health social networks: an observational study. J Med Internet Res 2014;16(2):e33 [FREE Full text] [doi: 10.2196/jmir.2966] [Medline: 24496109]
4. van Mierlo T, Hyatt D, Ching A. Mapping Power Law Distributions in Digital Health Social Networks: Methods, Interpretations, and Practical Implications. J Med Internet Res 2015;17(6):e160 [FREE Full text] [doi: 10.2196/jmir.4297] [Medline: 26111790]
5. Young C. Community management that works: how to build and sustain a thriving online health community. J Med Internet Res 2013;15(6):e119 [FREE Full text] [doi: 10.2196/jmir.2501] [Medline: 23759312]
6. Brownstein CA, Brownstein JS, Williams DS, Wicks P, Heywood JA. The power of social networking in medicine. Nat Biotechnol 2009 Oct;27(10):888-890. [doi: 10.1038/nbt1009-888] [Medline: 19816437]
7. Golder S, Donath J. Social roles in electronic communities. 2004 Presented at: Association of Internet Researchers (AoIR) Conference Internet Research 50; 2004; Brighton, England.
8. Wesler H, Cosley D, Kossinets G, Lin A, Dokshin F, Gay G. Finding social roles in Wikipedia. : ACM; 2011 Presented at: iConference; 2011; New York, NY.
9. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MJ. Self-reported differences in empowerment between lurkers and posters in online patient support groups. J Med Internet Res 2008;10(2):e18 [FREE Full text] [doi: 10.2196/jmir.992] [Medline: 18653442]
10. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
11. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
12. An LC, Schillo BA, Saul JE, Wendling AH, Klatt CM, Berg CJ, et al. Utilization of smoking cessation informational, interactive, and online community resources as predictors of abstinence: cohort study. J Med Internet Res 2008;10(5):e55 [FREE Full text] [doi: 10.2196/jmir.1018] [Medline: 19103587]
13. Børøsund E, Cvancarova M, Ekstedt M, Moore SM, Ruland CM. How user characteristics affect use patterns in web-based illness management support for patients with breast and prostate cancer. J Med Internet Res 2013;15(3):e34 [FREE Full text] [doi: 10.2196/jmir.2285] [Medline: 23454601]
14. Chomutare T, Arsand E, Fernandez-Luque L, Lauritzen J, Hartvigsen G. Inferring community structure in healthcare forums. An empirical study. Methods Inf Med 2013;52(2):160-167. [doi: 10.3414/ME12-02-0003] [Medline: 23392282]
15. Cobb NK, Graham AL, Abrams DB. Social network structure of a large online community for smoking cessation. Am J Public Health 2010 Jul;100(7):1282-1289 [FREE Full text] [doi: 10.2105/AJPH.2009.165449] [Medline: 20466971]

16. Myneni S, Cobb NK, Cohen T. Finding meaning in social media: content-based social network analysis of QuitNet to identify new opportunities for health promotion. Stud Health Technol Inform 2013;192:807-811. [Medline: 23920669]
17. Durant KT, McCray AT, Safran C. Social network analysis of an online melanoma discussion group. AMIA Jt Summits Transl Sci Proc 2010;2010:6-10 [FREE Full text] [Medline: 21347134]
18. Healey B, Hoek J, Edwards R. Posting behaviour patterns in an online smoking cessation social network: implications for intervention design and development. PLoS One 2014;9(9):e106603 [FREE Full text] [doi: 10.1371/journal.pone.0106603] [Medline: 25192174]
19. Jones R, Sharkey S, Smithson J, Ford T, Emmens T, Hewis E, et al. Using metrics to describe the participative stances of members within discussion forums. J Med Internet Res 2011;13(1):e3 [FREE Full text] [doi: 10.2196/jmir.1591] [Medline: 21239373]
20. Sudau F, Friede T, Grabowski J, Koschack J, Makedonski P, Himmel W. Sources of information and behavioral patterns in online health forums: observational study. J Med Internet Res 2014;16(1):e10 [FREE Full text] [doi: 10.2196/jmir.2875] [Medline: 24425598]
21. van Mierlo T, Voci S, Lee S, Fournier R, Selby P. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. J Med Internet Res 2012;14(3):e66 [FREE Full text] [doi: 10.2196/jmir.1854] [Medline: 22732103]
22. Zhao K, Greer G, Yen J, Mitra P, Portier K. Leader identification in an online health community for cancer survivors: a social network-based classification approach. Inf Syst E-Bus Manage 2014 Sep 16:1-17. [doi: 10.1007/s10257-014-0260-5]
23. Zhao K, Yen J, Greer G, Qiu B, Mitra P, Portier K. Finding influential users of online health communities: a new metric based on sentiment influence. J Am Med Inform Assoc 2014 Oct;21(e2):e212-e218 [FREE Full text] [doi: 10.1136/amiajnl-2013-002282] [Medline: 24449805]
24. Gruzd A, Haythornthwaite C. Enabling community through social media. J Med Internet Res 2013;15(10):e248 [FREE Full text] [doi: 10.2196/jmir.2796] [Medline: 24176835]
25. Stearns M, Nambiar S, Nikolaev A, Semenov A, McIntosh S. Towards Evaluating and Enhancing the Reach of Online Health Forums for Smoking Cessation. Netw Model Anal Health Inform Bioinform 2014 Dec;3:-. [doi: 10.1007/s13721-014-0069-7] [Medline: 26075158]
26. Chang H. Online supportive interactions: Using a network approach to examine communication patterns within a psychosis social support group in Taiwan. J Am Soc Inf Sci 2009 Jul;60(7):1504-1517. [doi: 10.1002/asi.21070]
27. Bambina A. Online Social Support: The Interplay of Social Networks and Computer-Mediated Communication. Youngstown, New York: Cambria Press; 2007.
28. Durant KT, McCray AT, Safran C. Identifying Temporal Changes and Topics that Promote Growth Within Online Communities: A Prospective Study of Six Online Cancer Forums. Int J Comput Models Algorithms Med 2011 Dec;2(2):1-22 [FREE Full text] [doi: 10.4018/jcmam.2011040101] [Medline: 22022636]
29. Wang X, Zhao K, Street N. Social Support and User Engagement in Online Health Communities. In: Smart Health.: Springer International Publishing; 2014 Presented at: International Conference for Smart Health; July 10-11, 2014; Beijing, China p. 97-110.
30. Borgatti S. Centrality and network flow. Social Networks 2005;27(1):55-71.
31. Freeman L. Centrality in social networks conceptual clarification. Social Networks 1978 Jan;1(3):215-239. [doi: 10.1016/0378-8733(78)90021-7]
32. Myeni S, Fujimoto K, Cobb N, Cohen T. Content-Driven Analysis of an Online Community for Smoking Cessation: Integration of Qualitative Techniques, Automated Text Analysis, and Affiliation Networks. Am J Pub Health 2015;105(6):1206-1212. [doi: 10.2105/AJPH.2014.302464]
33. MacQueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.: University of California Press; 1967 Presented at: The Fifth Berkeley Symposium on Mathematical Statistics and Probability; 1967; Berkeley, California.
34. Borgatti SP. Identifying sets of key players in a social network. Comput Math Organiz Theor 2006 Apr;12(1):21-34. [doi: 10.1007/s10588-006-7084-x]
35. Kleinberg JM. Authoritative sources in a hyperlinked environment. J ACM 1999;46(5):604-632. [doi: 10.1145/324133.324140]
36. Brin S, Page L. The anatomy of a large-scale hypertextual Web search engine. In: Computer Networks and ISDN Systems. 1998 Apr Presented at: Seventh International Conference on the World Wide Web; April 14-18, 1998; Brisbane, Australia p. 107-117. [doi: 10.1016/S0169-7552(98)00110-X]
37. Tausczik YR, Pennebaker JW. The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. J Lang Soc Psych 2009 Dec 08;29(1):24-54. [doi: 10.1177/0261927X09351676]
38. Wang Y, Kraut R, Levine J. Eliciting and receiving online support: using computer-aided content analysis to examine the dynamics of online social support. J Med Internet Res 2015;17(4):e99 [FREE Full text] [doi: 10.2196/jmir.3558] [Medline: 25896033]
39. Nunes S, Ribeiro C, David G. The impact of time in link-based Web ranking. Information Research 2013;18(3):paper 586.

40. Yu P, Li X, Liu B. On the temporal dimensions of search. In: Proceedings of the 13th International World Wide Web Conference. 2004 Presented at: 13th International World Wide Web Conference; 2004; New York, NY p. 448-449. [doi: 10.1145/1013367.1013519]

Abbreviations

HITS: hyperlink-induced topic search
IRRs: influential responding replies
OHC: online health community

Edited by G Eysenbach; submitted 17.05.15; peer-reviewed by S Myneni, M Ekstedt, E Børøsund, R Jones, K Zhao; comments to author 09.07.15; revised version received 26.08.15; accepted 07.10.15; published 01.12.15.

Please cite as:
Carron-Arthur B, Ali K, Cunningham JA, Griffiths KM
From Help-Seekers to Influential Users: A Systematic Review of Participation Styles in Online Health Communities
J Med Internet Res 2015;17(12):e271
URL: http://www.jmir.org/2015/12/e271/
doi:10.2196/jmir.4705
PMID:26627369

Original Paper

Online Outreach Services Among Men Who Use the Internet to Seek Sex With Other Men (MISM) in Ontario, Canada: An Online Survey

David J Brennan[1], MSW, PhD; Nathan J Lachowsky[1,2,3], PhD; Georgi Georgievski[1], HBSc; Brian R Simon Rosser[4], MPH, PhD; Duncan MacLachlan[5]; James Murray[6], MEd; Cruising Counts Research Team[1]

[1]University of Toronto, Factor-Inwentash Faculty of Social Work, University of Toronto, Toronto, ON, Canada
[2]University of British Columbia, Faculty of Medicine, Vancouver, BC, Canada
[3]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, BC, Canada
[4]University of Minnesota, School of Public Health, Minneapolis, MN, United States
[5]Ontario HIV Treatment Network, Toronto, ON, Canada
[6]Ontario Ministry of Health and Long Term Care, AIDS Bureau, Toronto, ON, Canada

Corresponding Author:
David J Brennan, MSW, PhD
University of Toronto
Factor-Inwentash Faculty of Social Work
University of Toronto
246 Bloor Street West
Toronto, ON, M5S1V4
Canada
Phone: 1 416 978 3273
Fax: 1 416 978 7072
Email: david.brennan@utoronto.ca

Abstract

Background: Men who use the Internet to seek sex with other men (MISM) are increasingly using the Internet to find sexual health information and to seek sexual partners, with some research suggesting HIV transmission is associated with sexual partnering online. Aiming to “meet men where they are at,” some AIDS service organizations (ASOs) deliver online outreach services via sociosexual Internet sites and mobile apps.

Objective: To investigate MISM's experiences and self-perceived impacts of online outreach.

Methods: From December 2013 to January 2014, MISM aged 16 years or older were recruited from Internet sites, mobile apps, and ASOs across Ontario to complete a 15-minute anonymous online questionnaire regarding their experience of online outreach. Demographic factors associated with encountering online outreach were assessed using backward-stepwise multivariable logistic regression ($P<.05$ was considered significant).

Results: Of 1830 MISM who completed the survey, 8.25% (151/1830) reported direct experience with online outreach services. Encountering online outreach was more likely for Aboriginal versus white MISM, MISM from Toronto compared with MISM from either Eastern or Southwestern Ontario, and MISM receiving any social assistance. MISM who experienced online outreach felt the service provider was friendly (130/141, 92.2%), easy to understand (122/140, 87.1%), helpful (115/139, 82.7%), prompt (107/143, 74.8%), and knowledgeable (92/134, 68.7%); half reported they received a useful referral (49/98, 50%). Few MISM felt the interaction was annoying (13/141, 9.2%) or confusing (18/142, 12.7%). As a result of their last online outreach encounter, MISM reported the following: better understanding of (88/147, 59.9%) and comfort with (75/147, 51.0%) their level of sexual risk; increased knowledge (71/147, 48.3%); and feeling less anxious (51/147, 34.7%), better connected (46/147, 31.3%), and more empowered (40/147, 27.2%). Behaviorally, they reported using condoms more frequently (48/147, 32.7%) and effectively (35/147, 23.8%); getting tested for HIV (43/125, 34.4%) or STIs (42/147, 28.6%); asking for their partners’ HIV statuses (37/147, 25.2%); and serosorting (26/147, 17.7%). Few MISM reported no changes (15/147, 10.2%) and most would use these services again (98/117, 83.8%). Most MISM who did not use online outreach said they did not need these services (1074/1559, 68.89%) or were unaware of them (496/1559, 31.82%).

Conclusions: This is the first online outreach evaluation study of MISM in Canada. Online outreach services are a relatively new and underdeveloped area of intervention, but are a promising health promotion strategy to provide service referrals and engage diverse groups of MISM in sexual health education.

(J Med Internet Res 2015;17(12):e277) doi:10.2196/jmir.4503

KEYWORDS

gay men; HIV prevention; Internet; online outreach; men who have sex with men; HIV/AIDS; sexual health; mobile technology and sexual health

Introduction

Background

Globally, the HIV epidemic continues to disproportionately burden sexual minority men. In Canada, this includes men who identify as gay, bisexual, two-spirit (used by the Aboriginal community to describe sexual minority individuals and/or nonbinary gender identity), as well as other men who have sex with men who may not necessarily identify as gay or bisexual—hereafter, all sexual minority men will be referred to as GB2M. In 2011 in Canada, there was an estimated 71,300 people living with HIV and nearly half of those were GB2M (n=33,330); nearly half of new HIV infections in 2011 were among GB2M (1480/3175, 46.61%) [1]. These figures have not changed much since 2008 [1]. As well, nearly one in five of all HIV-positive GB2M are unaware of their HIV infection [1]. Ontario—Canada's most populated province—accounts for the largest proportion (40.9%) of all HIV-positive tests reported nationally [2]. In Ontario, 60% of all people living with HIV/AIDS (PHA) are GB2M, who also accounted for 73% of all new diagnoses among men in 2012 [3].

Sexual Minority Men, HIV, and the Context of the Internet

GB2M have been using the Internet and online mobile technologies for well over 20 years to connect with one another for social and sexual relations [4]. Because of advancements in these technologies, men who use the Internet to seek sex with other men (MISM) have adapted to evolving technology that provides new options for connecting and obtaining sexual health information [4]. Even though many jurisdictions have advanced human rights for lesbian, gay, and bisexual people, many GB2M live, work, and socialize in contexts where same sex activity is stigmatized, and sexuality and sexual behavior are restricted or met with disapproval [5]; thus, many resort to online and mobile technologies to communicate and interact with other GB2M [6]. In a 2013 telephone interview study, 49.9% of GB2M in Canada reported using the Internet to look for sex in the past 6 months and 14% used a sociosexual mobile app (eg, Grindr) [7]. In another Canadian study published in 2013 but conducted online, 88% of participants used the Internet to find sex partners [8]. These two figures likely provide the bounds of the true estimate. Regardless, the Internet is the most frequently used resource for sex (eg, one-third of Ontario GB2M surveyed in 2006 used the Internet at least once a week to look for sex) [9].

In particular, these technologies offer what Cooper has referred to as the *Triple A Engine* effect [10]. That is, using online and mobile technologies for communication with other MISM for social and sexual reasons is appealing and common because of three factors: Affordability, Accessibility, and Anonymity. Subsequent additions to this model include Acceptability (greater tolerance online) [11], Approximation (greater ability to experiment) [12], and most recently Affirmation (explore and confirm one's identity) and Assessment (ability to prescreen and assess compatibility of potential partners) [13]. The components of this descriptive framework can assist researchers and providers in understanding how online outreach may be an important and beneficial tool to address online HIV/sexually transmitted infection (STI) prevention and sexual health promotion among MISM.

Online Partners and Sexual Behavior

GB2M use electronic media to look for sex, friendship, and connection online. GB2M who seek sexual partners online report high rates of behaviors associated with HIV risk (eg, unprotected anal intercourse [UAI] with a serodiscordant partner) [14]. Among a 2008 community-based sample of GB2M in British Columbia, Canada, men who sought partners online were more likely to report 10 or more sexual partners in the past year compared with those who did not seek partners online [15].

Online HIV/Sexually Transmitted Infection Prevention

Some AIDS service organizations (ASOs), community-based organizations (CBOs), public health groups, and Internet providers have developed programs and models that work directly with MISM to support sexual health. MISM appear to lack basic knowledge of HIV (eg, how HIV is transmitted and how condoms should be used), have had questions about HIV testing, and feel that community resources do not meet their needs [16]. Previous research has shown that the majority of MISM hold favorable attitudes toward online health promotion [14,17-19]. Online sexual health promotion can be an effective and low-cost method to educate MISM [18,20,21]. A 2013 study of MISM aged 18-24 years in Southern California found that the number one reason for using Grindr, a sociosexual mobile app, was to meet hookups (for sex); 70% of young MISM expressed a willingness to participate in app-based HIV prevention [22]. Online interventions have demonstrated a reduction in UAI [23], particularly among unknown or serodiscordant partners [24]. Other research has demonstrated efficacy in online interventions to increase HIV/AIDS knowledge, self-efficacy, and condom use among MISM [25]. As participants, MISM expect online health promotion to respect the online culture, build trust, and deliver well-crafted and focused messages [16]. A recent systematic literature review [26] on Internet and mobile app use for sexual health promotion among MISM in Canada highlighted the need for more research that examines this phenomenon and its implications. Specifically, though agencies are offering online outreach services, we know very little about how these online outreach services are understood, accessed, or relevant for the sexual health of MISM in Canada.

Online Outreach to Men Who Use the Internet to Seek Sex With Other Men in Ontario, Canada

The ASO and sexual health sectors, as well as public health providers and public health scientists, have suggested a need to reconceptualize online media as vital tools for HIV prevention [4,27]. Though the Internet has been used in a number of ways for outreach to MISM (eg, e-blasts, social media, and chat rooms) over the past two decades, this study is focused on the notion of online outreach services that involve trained staff and/or volunteers logging into online sites and apps and responding to questions from, and providing referrals for services to, MISM in these environments. For the purposes of this research, online outreach is broadly defined as the delivery of information and/or support services regarding HIV, STIs, and general health with a particular focus on sexual health via any Internet website, mobile phone app, or Web-based tool that MISM use as a means of connecting with other MISM for social/sexual activity. Resources and geography limit comprehensive online outreach for MISM in Ontario. In terms of geographic coverage, while there is some Internet-based service provision in all regions of Ontario, 70% of online outreach is provided by agencies in Toronto, the province's most populous metropolitan area [3]. It is important to note that for resource reasons, most of this online outreach is conducted at varying times of the day, including weekends, weekdays, and weeknights, depending on the provider's capacity and availability. It is not a 24-hour service and it is not available through all apps or websites. Some app or website providers prohibit this type of service on their sites as it interrupts the user experience. In Ontario, online outreach providers are generally either trained sexual health outreach workers or public health staff.

Online outreach has become a key tool in the delivery of sexual health information and services affecting MISM [3,8,22,28,29]. In light of the predominance of the Internet as a social and sexual venue for GB2M in Ontario, outreach in physical venues is no longer sufficient; online outreach can help reach more GB2M. This study developed out of consultations with community providers who were conducting online outreach to MISM. Often these providers were doing this outreach because they were aware that this is where the men they wished to reach were located, or because physical venues were not available or less popular. Given the numerous calls for more research directed at understanding how online tools can benefit sexual health promotion [26,30], few of these articles focused specifically on online outreach. Therefore, there is little evidence of the reach, depth, impact, or effectiveness of this type of outreach. Though ASOs and public health practitioners have developed apps specific to HIV prevention, these are infrequently downloaded and often poorly rated, suggesting that these apps may not be useful or that MISM prefer accessing information within the apps they use [31]. Agencies in Ontario have reported that "there is still some uncertainty about how to do Internet-based outreach well, as well as ongoing challenges in tracking and assessing the impact of this work" [3]. Several agencies report an increased demand for online outreach services [3]. Taken collectively, there is great interest in improving online outreach, but also some challenges conceptualizing, sustaining, and evaluating online media as effective learning environments.

Without better evidence and understanding of how MISM are seeking and/or experiencing online outreach, there remains a missed opportunity to connect with GB2M "where they are at" with the goals of preventing HIV/STI transmission and improving sexual health. This study sought to examine how MISM in Ontario access, experience, and perceive online outreach. That is, whether they find it useful, relevant, and applicable to their sexual health. The aim of this study was to assist in the development of tools that would be useful for sexual health online outreach for MISM.

Methods

Study and Participants

Data were drawn from a mixed-methods, community-based research study entitled *Cruising Counts*, which involved partnerships from across Ontario, Canada's largest province. The research team included various frontline staff and managers from ASOs who were providing or had provided online outreach services to MISM, staff of the provincial Gay Men's Sexual Health Alliance, a provincial HIV/AIDS health policy expert, and researchers from three academic institutions. A community advisory board composed of MISM met quarterly to inform and provide feedback on the research process, data collection tools, and knowledge translation activities. All study protocols were granted ethics clearance from the University of Toronto Research Ethics Board.

Between December 2013 and January 2014, participants were recruited from across Ontario using electronic advertisements on sociosexual websites (eg, Squirt.org, recon.com, bgclive.com, and craigslist.ca), mobile apps (eg, Grindr), social media (eg, Facebook and Twitter), and printed flyers distributed through ASOs. Participants were asked to complete an anonymous online questionnaire regarding their technology use, online behavior (sociosexual and health related), experience of online outreach, and demographics (see Table 1). To be eligible, participants must have sought sexual partners or sexual health information online in the past 6 months (or had an interest in doing so); identify as a gay, bisexual, two-spirit, straight, queer, or questioning cis- or transgender man; had in the past had sex with another man (or an interest in doing so); be at least 16 years old; and either lived or worked in Ontario or had visited Ontario at least three times in the past year. Participants were offered an opportunity to enter a random draw for one of 40 cash prizes of Can $100, which were delivered via Interac e-Transfer.

Measures

Online Outreach

Our primary dependent variable was participants having experienced online outreach or not. Participants were prompted with the following definition: “By online outreach services, we mean that while you were online or using an app, you had any interactive conversations, whether in real time or not (can include chatting, responding to postings/message boards, or messaging), between you and an online outreach worker.” Participants were asked, “Have you ever encountered or used online outreach services?” and to indicate who initiated the contact (participant or provider). Those who were contacted by online outreach services were asked if they were told why they were being contacted, whether a photo was used (agency logo, personal, unsure), and whether they were comfortable being contacted.

All participants who had experienced online outreach rated several aspects of their last experience (see Table 2 for items) on a 5-point Likert scale from 1 (disagree) to 5 (agree), which were dichotomized into agree (4 or 5) or not (1-3). Participants were asked to indicate any changes that resulted from their last online outreach encounter (see Table 3 for items), if they would use online outreach services again (yes or no), and to explain why or why not through an open-ended text response. Those participants who had not encountered online outreach were asked to indicate why: *no need/interest*, *not available when needed*, *don’t trust*, or *don’t know about it or where to find it*. All participants were asked to provide qualitative feedback on any difficulties they had trying to access these online outreach services if these services were of interest to them.

Demographics

Independent variables for this analysis included age (in years), race/ethnicity (white, black, Latino, Aboriginal, South Asian, Southeast/East Asian, mixed race, or other), sexual orientation (gay/homosexual, bisexual, or other), gender (cis-gender male or not), two-spirit status (yes or no), trans person (yes or no), student (yes or no), social assistance (Ontario Disability Support Program [ODSP]/Employment Insurance [EI]/Ontario Works or not), highest level of formal education attained (no postsecondary education, some postsecondary education, or finished postsecondary education or any postgraduate education), immigration status (Canadian citizen/permanent resident or not), and self-reported HIV status (HIV negative, HIV positive, or unknown). Further, geographic location was grouped into major provincial regions (Toronto, Central Ontario, Southwestern Ontario, Eastern Ontario, and Northern Ontario) using participants' forward sortation areas (ie, first three characters of a Canadian postal code).

Analyses

All quantitative data analyses were conducted using the statistical package Stata/SE version 13 (StataCorp) and $P<.05$ was considered significant unless otherwise specified. Data were analyzed to determine the prevalence of online outreach experience and associated factors. Descriptive statistics of the overall sample and for those participants who experienced online outreach were prepared. Independent factors associated with experience of online outreach were determined using logistic regression. Univariate analyses were conducted to screen independent variables using a liberal P value of .20 [32]. A final multivariate model was built using a manual backward-stepwise elimination approach [32]. Nonsignificant likelihood ratio tests were used to confirm removal of any categorical variables. Confounding was manually assessed throughout model building; if the addition or removal of a variable resulted in a greater than 30% change in any other independent variable’s coefficient, it was retained in the model [32]. A research assistant manually coded qualitative data, which were collected through open-ended text responses, with iterative consultations with the first and second author (DJB, NJL) and to resolve unclear coding questions.

Results

Of the 1830 men who completed the online questionnaire, 151 men (8.25%) reported experience with online outreach, 95 men (5.19%) were unsure if they had, and 25 men (1.37%) refused to answer the question. Table 1 provides the descriptive statistics of the overall sample and the prevalence of, and univariate associations with, reporting online outreach experience. Two factors that were significant at the univariate level, but that were not included in the multivariable analysis, were HIV status and being two-spirited. Compared with HIV-negative men, HIV-positive men were more likely to have experienced online outreach (odds ratio [OR] 2.19, 95% CI 1.35-3.55). Also, two-spirit participants were more likely to have also experienced online outreach compared with those who were not two-spirit (OR 3.38, 95% CI 1.33-8.58). The final multivariate model included race/ethnicity, location, and social assistance as independent factors associated with having experienced online outreach (see right-most column in Table 1). Aboriginal men were more likely than white men to have experienced online outreach (adjusted odds ratio [AOR] 2.75, 95% CI 1.03-7.29). Compared with men who lived in Toronto, men in Southwestern Ontario (AOR 0.49, 95% CI 0.28-0.84) and Eastern Ontario (AOR 0.60, 95% CI 0.37-0.97) were less likely to have experienced online outreach. Finally, men who were receiving some form of social assistance (eg, disability or unemployment insurance) were more likely to have experienced online outreach than those who were not receiving social assistance (AOR 3.23, 95% CI 1.96-5.31).

Table 1. Sample demographics and the prevalence of, and factors associated with, online outreach experience.[a]

Demographics	Overall sample (n=1830), mean (SD) or n (%)	Experienced online outreach (n=151), mean (SD) or n (%)	Univariate associations, OR[b] (95% CI)	Multivariate associations, AOR[c] (95% CI)
Age in years, mean (SD)	37.8 (13.2)	36.6 (13.1)	0.99 (0.98-1.01)	Not included
Race/ethnicity, n (%)				
White	1448 (79.13)	117 (78.5)	1.00	1.00
Black	33 (1.80)	5 (3.4)	2.09 (0.79-5.52)	1.24 (0.41-3.73)
Latino	45 (2.46)	4 (2.7)	1.13 (0.40-3.20)	0.95 (0.32-2.81)
Aboriginal	27 (1.48)	6 (4.0)	3.38 (1.33-8.58)[d]	2.75 (1.03-7.29)[d]
Other	29 (1.58)	2 (1.3)	0.90 (0.21-3.85)	0.87 (0.20-3.79)
South Asian	40 (2.24)	2 (1.3)	0.59 (0.14-2.49)	0.58 (0.14-2.48)
Southeast/East Asian	112 (6.12)	6 (4.0)	0.64 (0.28-1.50)	0.53 (0.21-1.36)
Mixed race	70 (3.83)	7 (4.6)	1.27 (0.57-2.84)	1.14 (0.50-2.57)
Sexual orientation, n (%)				
Gay	1325 (72.40)	117 (77.5)	1.00	Not selected
Bisexual	438 (23.93)	33 (21.9)	0.85 (0.57-1.26)	
Other	63 (3.44)	1 (0.7)	0.16 (0.02-1.20)	
Two-spirit, n (%)				
No	1789 (97.76)	144 (95.4)	1.00	Not selected
Yes	36 (1.97)	7 (4.6)	2.72 (1.17-6.32)[d]	
Trans, n (%)				
No	1800 (98.36)	147 (97.4)	1.00	Not selected
Yes	25 (1.37)	4 (2.7)	2.11 (0.72-6.24)	
Cis-gender male, n (%)				
No	31 (0.05)	5 (3.3)	1.00	Not selected
Yes	1795 (98.09)	146 (96.7)	0.47 (0.18-1.23)	
HIV status, n (%)				
HIV positive	146 (7.98)	23 (15.4)	2.19 (1.35-3.55)[d]	Not selected
HIV negative	1439 (78.63)	113 (75.8)	1.00	
Unsure	217 (11.86)	13 (8.7)	0.74 (0.41-1.34)	
Location, n (%)				
Toronto	512 (27.98)	55 (37.9)	1.00	1.00
Central Ontario	315 (17.21)	24 (16.6)	0.69 (0.42-1.13)	0.67 (0.40-1.13)
Southwestern Ontario	350 (19.13)	21 (14.8)	0.53 (0.31-0.89)[d]	0.49 (0.28-0.84)[d]
Eastern Ontario	421 (23.00)	29 (20.0)	0.61 (0.38-0.98)[d]	0.60 (0.37-0.97)[d]
Northern Ontario	121 (6.61)	16 (11.0)	1.27 (0.70-2.30)	1.03 (0.55-1.95)
Education, n (%)				
High school or less	248 (13.55)	24 (15.9)	1.38 (0.83-2.29)	Not included
Some postsecondary	855 (46.72)	74 (49.0)	1.18 (0.82-1.71)	
Bachelor’s degree or greater	713 (38.96)	53 (35.1)	1.00	
Student, n (%)				
No	1505 (82.24)	127 (85.2)	1.00	Not included

Demographics		Overall sample (n=1830), mean (SD) or n (%)	Experienced online outreach (n=151), mean (SD) or n (%)	Univariate associations, OR[b] (95% CI)	Multivariate associations, AOR[c] (95% CI)
	Yes	295 (16.12)	22 (14.8)	0.87 (0.54-1.39)	
Social assistance, n (%)					
	No	1680 (91.80)	124 (83.2)	1.00	1.00
	Yes	120 (6.56)	25 (16.8)	3.37 (2.09-5.44)[d]	3.23 (1.96-5.31)[d]
Canadian citizen/permanent resident, n (%)					
	No	82 (4.48)	6 (4.0)	1.00	Not included
	Yes	1748 (95.52)	145 (96.0)	1.16 (0.50-2.72)	

[a] Missing values excluded from this table.

[b] OR: odds ratio.

[c] AOR: adjusted odds ratio.

[d] *P*<.05.

Most men had contacted the service provider directly (125/151, 82.8%). For those who had been contacted by an online outreach worker (26/151, 17.2%), most participants reported that the worker explained why they were contacting them (16/26, 62%), that the worker had an agency logo (12/26, 46%) or photo of themselves (5/26, 19%), and that they were comfortable being contacted (17/26, 65%).

Participants rated their last online outreach experiences very positively (see Table 2). Men reported that the individual was friendly (130/141, 92.2%), used language they could understand (122/140, 87.1%), was helpful (115/139, 82.7%), was prompt to reply (107/143, 74.8%), and was knowledgeable and a trusted source of information (92/134, 68.7%). Over three-quarters of men were comfortable (116/144, 80.6%) and felt satisfied (110/142, 77.5%) with the interaction. Half of the men were provided with a useful referral (49/98, 50%). Very few men rated their last online outreach worker as confusing (18/142, 12.7%) or invasive or annoying (13/141, 9.2%).

Table 2. Participants' ratings of their last online outreach encounter.[a]

Survey items	Agreed, n (%)
The individual was friendly (n=141)	130 (92.2)
The individual used language I could understand (n=140)	122 (87.1)
The individual was helpful (n=139)	115 (82.7)
I was comfortable with the interaction (n=144)	116 (80.6)
I was satisfied with the interaction (n=142)	110 (77.5)
The individual was prompt to reply (n=143)	107 (74.8)
The individual was knowledgeable and a trusted source of information (n=134)	92 (68.7)
The individual provided me with a useful referral (n=98)	49 (50)
The individual was confusing (n=142)	18 (12.7)
The individual was invasive or annoying (n=141)	13 (9.2)

[a] Missing values excluded from this table.

Table 3 shows the number and proportion of men who self-reported a variety of impacts as a result of their last online outreach experience. Over half of the men reported a better understanding of (88/147, 59.9%), and an increased comfort about (75/147, 51.0%), their level of sexual risk. Online outreach connected some men with a variety of referral services for men: 34.4% (43/125) got an HIV test, 28.6% (42/147) got an STI test, 12.2% (18/147) sought out counseling, 9% (2/22) of HIV-positive men sought HIV-related care, and 6.8% (10/147) of men got STI treatment. More frequent (48/147, 32.7%) and effective (35/147, 23.8%) use of condoms was also reported.

Men also reported changes in their sexual partnering decisions with respect to HIV status: 25.2% (37/147) reported only having sex with partners whose HIV status they knew, 17.7% (26/147) reported only having sex with seroconcordant partners, and 4.8% (7/147) reported only having sex with HIV-positive partners whose viral load they knew. Online outreach also seemed to benefit men's social, mental, and emotional well-being; some men felt less anxious (51/147, 34.7%) and others felt better connected (46/147, 31.3%), more empowered (40/147, 27.2%), and more sexually satisfied (20/147, 13.6%).

Table 3. Self-reported impact as a result of last online outreach encounter (n=147).[a]

Survey items	n (%)
I better understand my sexual risks	88 (59.9)
I am more comfortable about my level of sexual risks	75 (51.0)
I increased my knowledge	71 (48.3)
I feel less anxious	51 (34.7)
I got an HIV test (only for HIV-negative or status unknown men, n=125)	43 (34.4)
I use condoms more frequently	48 (32.7)
I feel better connected	46 (31.3)
I got an STI[b] test	42 (28.6)
I feel more empowered	40 (27.2)
I made decisions to	
...only have sex with people whose HIV status I knew	37 (25.2)
...only have sex with people who had the same HIV status as I do	26 (17.7)
...only have sex with HIV-positive people whose HIV viral load I knew	7 (4.8)
I use condoms more effectively (without slips, tears, or breakage)	35 (23.8)
I feel more sexually satisfied	20 (13.6)
I sought out counseling	18 (12.2)
I made no changes	15 (10.2)
I got HIV care (only for HIV-positive men, n=22)	2 (9)
I got STI treatment	10 (6.8)

[a] Missing values excluded from this table.

[b] STI: sexually transmitted infection.

The vast majority of participants (98/117, 83.8%; 34 refused to answer) who used online outreach services said that they would use them again. When asked to explain why qualitatively, 86 out of the 98 (88%) men provided reasons that were thematically coded; convenience (24/86, 28%), reliability (22/86, 26%), and anonymity (20/86, 23%) were the most commonly cited reasons for future use. For example, one man stated that these services were “available when needed (24 hrs) and voluntary,” while another explained that “the information was excellent and private.” Others reported they appreciated that these services offer “someone understanding, nonjudgmental, and open to discussing my concerns and answering my questions,” “immediate contact with a compassionate person,” and “human contact.” One man remarked supportively, “I find it difficult to ask the same questions with health care providers face-to-face because I have had negative, homophobic experiences in the past.”

The small minority of men who had previously used online outreach services, but indicated that they would not use them again (19/117, 16.2%), expressed three main reasons for this: (1) negative experiences or perceptions of these services, (2) long wait times to get a response, and (3) a preference or opportunity to interact with health professionals in person. One man stated that the experience “was invasive and [I] couldn’t feel as though I could trust them.” With regard to responsiveness, participants explained that they were “too slow to answer,” “received no response,” and that “the first time I tried contacting someone [in] real time they never replied back.” Finally, some of these men spoke positively about their current access to gay-friendly health services: “It was easier to just go to the nearest health clinic” and “I have an excellent support system with my doctor already.”

Overall, the men who accessed online outreach services rated the services as helpful and several explained how these services were becoming important sources of community and service information. For example, one man explained, “I find it difficult sometimes to remember when drop-in anonymous testing hours/locations happen, and the reminder that there is one in my 'hood is nice.” Participants provided a number of recommendations to consider in the future provision or adaptation of online outreach services. First, even among those who had accessed online outreach services, several men expressed that these services were not readily visible or available. One man stated that these services “need [a] more visible presence online” and another explained, “The only downside is that the individual has to seek out and discover the services available to them—so having this sort of thing more accessible on apps...would be a good idea.” Several men expressed an interest in more real-time conversations: “I wish the interaction could be simultaneous. My responses often had a long delay (like a day or two).” Others explained this in terms of better geographic, temporal, and online venue coverage: “They were helpful with answer[ing] my questions but...they could not refer me to local ones,” “more outreach volunteers in different locations," and "They need to be on the sites that we are on, and to be there round the clock 24/7."

Participants who had explicitly not experienced online outreach (n=1559) said they had not used these services because they had no need or interest (1074/1559, 68.89%), did not know about or where to find them (496/1559, 31.82%), did not trust them (85/1559, 5.45%), or the services were not available when they needed them (76/1559, 4.87%). The few respondents who said they did not trust these services explained that "I'm somewhat intimidated by the lack of privacy," "Trust would be an issue, as well what personal information would be needed before the advice is given," and they "need complete secrecy." Many of the individuals who said that they did not need or were not interested in these services explained that "I go to the clinic about once a year and talk with nurses" or "I prefer to talk with a professional in person!"

All participants were asked to comment on any difficulties they had in trying to access online outreach services. Of 1830 men, 1005 (54.92%) were not interested in accessing these types of services; a further 412 (22.51%) had never experienced any barriers to accessing these services. Of those men who had not accessed online outreach and indicated that they wished to, but had potential difficulties, half of the reasons were either not knowing how to access these services (75/296, 25.3%) or being unaware that these services existed prior to this questionnaire (67/296, 22.6%). Participants reported that they "did not know they [online outreach services] existed or where to find them—I really never heard of this before, otherwise I would have contacted them" and "I had no idea these services existed...so more promotion of these services in the gay community would have been nice."

Others indicated geographic difficulties with accessing these services. Several participants from smaller cities explained this poses almost paradoxical problems where gay-friendly services are not available (eg, "No access near where I live"), but attempts to make them available pose particular challenges to anonymity. One man explained that he lived in a "small city that hires mostly LGBT [lesbian, gay, bisexual, and transgender] staff. This will not help accessing services as who wants to run into a friend or acquaintance when seeking assistance?" Another participant stated, "In my small town, I still find that looking [for] and getting sexual health info or treatment has a huge stigma attached to it. Traveling to [Toronto] is costly and I can't do it easily." Access to Toronto-based services was also a challenge described for men comparably closer: "Services seem to be located downtown Toronto and I live in the suburbs." Other geographic barriers were related to technology access and coverage: "limited Internet service where I live" and "Unless you're located in the core of Toronto [major metropolitan center], it's likely you won't see an outreach worker on Grindr." Some participants expressed that they "don't know how to tell if info is reliable," are concerned with "having to give personal information...and feeling like you're being judged," or that the "gay outreach community is too small and not sure if what I share will remain confidential."

Discussion

Principal Findings

This is the first study to examine the perceptions, expressed need, and self-reported impact of online outreach services for MISM (GB2M) in a Canadian context. Similar to findings in the United States [33], our study suggests that online outreach is a useful and important tool in HIV prevention for GB2M. The findings suggest that 8% of MISM in this sample have accessed online outreach services. The community members of the research team considered the reaching of 8% of all GB2M (MISM) through online outreach in their communities to be very successful, since the size of the population online is so large and since they represent a handful of agencies that have been offering these services only in the past couple of years. It is significant that participants are encountering these services in a space online where they are not necessarily going to seek health information and services. Most importantly, our analyses showed online services to be disproportionately used by GB2M who are hard to reach using other means, including HIV-positive men and Aboriginal men. Though several other studies focused primarily on youth [22,25], our findings suggest that there were no differences based on age. It would be useful to ensure broad age ranges for studies examining GB2M and online use.

In univariate analyses, two-spirit men and HIV-positive men were significantly more likely to access online outreach, and for trans-identified men this analysis approached significance. However, in multivariable analysis, Aboriginal participants (not necessarily those who identified as two-spirit) were more likely to access online outreach compared with white men. This may be a useful distinction for agencies serving Aboriginal men, because regardless of how they identify—gay, bisexual, two-spirit—this suggests Aboriginal men of all identities—gay, bisexual, two-spirit—are more likely to access resources available online. Like non-Aboriginal communities, stigma among Aboriginal populations regarding sexual minorities [34] may leave some two-spirit men feeling that online outreach is the preferred place to encounter the Triple A Engine effect of service usage because it is accessible, affordable, and anonymous. Those in Southwestern and Eastern Ontario were less likely than those in Toronto to access outreach. This may be a result of the numerous agencies providing online outreach in Toronto, whereas smaller communities may likely have only one (if any) agency providing such outreach. However, there were no differences in online outreach uptake between men in Northern Ontario, which has more rural and remote regions, and men in Toronto; this indicates the utility of these services to reach men who may be more geographically isolated [25] from both physical communities and in-person health and social services that are often clustered within large urban centers.

Those on social assistance were more likely to access online outreach than those who were not. This may speak to the affordability and accessibility of online resources and the lack of physical barriers/challenges to access them. These factors are important considerations for programming and policy implications when developing strategies in a variety of local jurisdictions and with specific populations. Indeed, a diverse array of singular and multiple outcomes were noted, including a decrease in anxiety, an increase in condom use, and even a better connection to community. These outcomes suggest online outreach has the capacity to address a comprehensive and integrated approach to care and services. Clearly, the needs go beyond specific condom use issues to those involving access to testing, pre- and post-HIV exposure treatment, and issues related to mental health and social well-being. There is a growing consensus that to address sexual health among GB2M, providers must understand the linkages between HIV/STIs, other health issues, and the social determinants of health along with the notion that multiple health epidemics—substance use, childhood trauma and bullying, mental health—are part of a fabric of syndemics that impact the health and wellness of GB2M across the life spectrum. These findings suggest the use of the Internet and other mobile technologies is a yet unrealized potential tool to link GB2M to rapidly evolving HIV-prevention information, such as the meaning of an undetectable viral load in relation to HIV transmission, and care, such as pre- and postexposure treatment. In addition, these online tools can also be formulated to address the underlying psychosocial factors impacting syndemics among GB2M [27]. Our findings support previous calls to encourage a more holistic approach to the health of MISM [4,27].

Of the men who accessed online outreach, their experiences were mostly favorable, finding the contact helpful and relevant, and reporting that they would use the service again. Half reported that they increased their knowledge on sexual health. Many men reported that they received a referral for testing or other services as an outcome. Nearly 90% of men who reported connection with online outreach reported some change as a result of that interaction. A recent systematic review of the literature has shown that using online tools for HIV prevention among gay and bisexual men were very effective at creating behavior change related to HIV transmission among gay and bisexual men [35].

Of the small number who would not use online outreach again, they primarily reported that they had a negative experience, a long wait time, or a preference for in-person contact. For those who did not access online outreach services, they reported mostly that they did not have a need or interest in the service, they did not know the service existed, or the services were not available at the time required. This feedback is useful when considering the development and implementation of online outreach services. Some regions may be lacking in online outreach due to resource constraints. In addition, MISM may be on certain apps or sites that have little or no online outreach and thus may not be coming into contact with such services. In our sample, most men reported being on numerous sites/apps (data not shown). Of course, it is not feasible that online outreach services can be provided on all sites or apps at all times, but, similar to previous research [28], it appears that MISM are interested in these services and a greater saturation of services may be helpful. It may be beneficial to have a sense of where men are appearing online and attempt to target the most popular sites/apps for a particular population or location, although these also shift over time as new products are released on the market. Therefore, it is important that agencies be aware of these changes in their population and are able to transition/adapt service provision across various platforms, including new and emerging ones [31].

Some men preferred the anonymity of the Internet for health care resources while others felt they had a trusting relationship with their providers. Thus, having resources online and on apps may be an important adjuvant to in-person care [28]. These services cannot and should not be a replacement for in-person care, but it seems clear there is a high demand and interest in online outreach, and that these services are commonly used by diverse MISM who may experience barriers to traditional services [8].

A large proportion of men reported not accessing these services due to a lack of need or interest (1074/1559, 68.89%). Future research would do well to understand the ways in which these men feel they have their health needs met to see if their resources, skills, or knowledge are transferable to other men. They may be using other forms of outreach, have advanced health literacy, or have access to other pertinent health information and care. It may also be that these men may have heightened HIV/STI risk, but do not perceive themselves to be at risk.

Limitations

These findings were primarily self-reported and therefore may be impacted by recall and response bias. Participants were recruited through online venues and agency email blasts and therefore may not accurately capture the characteristics or number of GB2M who are online or using online outreach services. Our findings may not be generalizable to MISM populations across Ontario or in other jurisdictions, but they do provide an indication of important trends that should be investigated in population-based studies and with different study designs. In addition, the small sample size of men who accessed services may mask some of the other differences. Though there were location, race, and socioeconomic differences in those accessing online outreach, our study design could not help us to understand what the reason for these differences were. Finally, the scope of our conclusions is also limited by the lack of information related to funding history, reasons for starting and ending online outreach programs, and issues of program sustainability.

Comparisons With Prior Work

Similar to previous research [36], our findings suggest that MISM are actively using the Internet for information regarding HIV and other STIs. Given that previous research has suggested that GB2M in online environments may lack basic HIV education [17], this study's findings resonate with previous research that suggests that GB2M are increasingly using the Internet and mobile technologies for sexual contact and sexual health promotion and information. Our findings also show that GB2M are indeed open and willing to use a variety of online resources to access HIV/STI and other health information [4,33]. Previous research has reported this acceptability in a variety of contexts, specifically online forums [18], Web tools [24], mobile-based apps [22], social media networks (ie, Facebook and YouTube) [37], and highly interactive online virtual environments [23]. The authors attempted to compare the levels of interaction and acceptability of online outreach to other forms of outreach among other GB2M. Though there is robust research focused on the outcomes and use of other forms of outreach, such as bar and bathhouse testing and counseling, there is a paucity of recent research that evaluates the acceptability levels of these tools among GB2M, thus limiting any possibility to compare. However, evidence suggests that previously developed mobile apps aimed at reducing HIV and other STIs have not been rated very well and were not frequently downloaded [31]. It is possible that instead of developing separate apps, engagement in online outreach within currently used apps may be more acceptable to MISM. Future research is needed to examine this possibility. Future research is also needed to test the efficacy and effectiveness of online outreach as a health intervention using experimental or quasi-experimental designs. An economic analysis of online outreach as a health service could help demonstrate the breadth of potential positive outcomes that occur. Additional useful research would be to examine the ways in which those who did not need or desire online outreach understand their sexual health and what they do to maintain it. Comparative research aimed at testing the levels of acceptability to various forms of online outreach for GB2M would provide beneficial data to ascertain the relevance and need for online versus more traditional forms of outreach (eg, bars and bathhouses).

Previous research has also suggested that rural GB2M are willing and interested in online tools for HIV/STI prevention and that such interventions can be efficacious as well [25]. Given Ontario’s large rural areas, it would be important for future research to examine the particular needs and usefulness of online outreach for rural and nonmetropolitan MISM. Our findings suggest that some GB2M desire to have resources accessible in person. This finding resonates with Hottes et al [31] who found that online testing was unlikely to replace in-person HIV testing among gay men, but may be a useful option for some who lack access to resources that are knowledgeable about gay men’s sexual health.

From a theoretical perspective, we found that Cooper’s Triple A Engine effect [10] and its offshoot components [11-13] were especially relevant to the experience of MISM using online outreach services. The elements of Cooper’s work and others who have developed it further [10-13] generally correspond with the data found in our study. These include that online outreach services are affordable, accessible (if one has a mobile phone or Internet connection), and anonymous. Though participation in the study required Internet access, any outreach experienced would be anonymous and free (affordable). In terms of accessibility of the services themselves for those online, the main barrier was a lack of awareness of the service or how to find it. This can help providers to consider better ways to increase awareness of their services. Other components related to the Triple A Engine effect suggested by other authors include Acceptability, Affirmation, and Assessment. Acceptability [11] was moderately evident in this study and future research should more directly examine whether having online outreach in apps is acceptable to GB2M; other research has shown that GB2M are interested in using online tools and resources for sexual health [22,28,31]. Affirmation [13] of oneself and a connection to a community were reported by those who felt they could trust the community providers. Assessment [13] of one’s sexual risk and making changes to one’s behavior as a result of this outreach was reported by most of the men who had encountered online outreach. Clearly, the use of online outreach represents an important and emerging tool to support the sexual health of MISM.

Conclusions

Online outreach for HIV/STI prevention is a promising tool for GB2M and the agencies that serve them. For GB2M who access them, these services are helpful, provide useful referrals, and appear to provide some self-reported change in knowledge and behavior. For those who did not use online outreach, the primary reason was that they felt it was not necessary. For some men the response time was too slow or the services were challenging to find. These responses suggest providers might be interested in ensuring shorter response times and more awareness campaigns that inform GB2M about the scope and availability of their online and Web-based services. Shorter response times are important as GB2M who may be exposed to HIV may want to access postexposure prophylaxis (PEP) treatment, which should begin as soon as possible (within 72 hours) of exposure.

These findings suggest that online outreach services show great promise to reach some hard-to-reach populations. Further research may be needed to understand why men from certain regions, or those who are Aboriginal or on social assistance, may use online outreach services more. It is possible that those with marginalized status may have fewer local resources available to them and therefore using the Internet for information and referrals is more convenient and accessible. As per the expanded Triple A Engine effect concepts, because online outreach services are typically free and anonymous, this may be driving the increased usage among those with lesser financial resources and those who are closeted or more likely to suffer stigma. These findings also suggest opportunities for funders and service providers of online outreach to increase the reach and the uptake of such services to reduce the impact and burden of HIV and other STIs among GB2M. Additionally, because various mobile apps and websites cater to specific populations, online outreach has the advantage of providing more population-specific and individualized responses. The outreach can be provided in such a way as to specifically address one person’s current needs within an understood culture. Examples of this include Aboriginal men, men of color, rural men, younger men, and older men.

Finally, these findings suggest that future HIV-prevention interventions aimed at GB2M consider the further development and coverage of online outreach programs and services. Given the positive appreciation of this type of outreach and the accessibility and anonymity it provides, online outreach is an important and emerging tool that has the potential to address the broad range of issues that fuel the HIV epidemic among GB2M; these issues include the dissemination of accurate and timely information, as well as access to testing, care, and services that address the broad range of psychosocial issues that impact GB2M.

Acknowledgments

We would like to thank all of our participants for sharing their important data. The *Cruising Counts* study was funded by the Ontario HIV Treatment Network 2013 Targeted Request for Applications (RFA: G913 & RFA: G847). DJB is supported by an Ontario HIV Treatment Network Applied HIV Research Chair. NJL was supported for this work by a Canadian Institutes of Health Research Vanier Canada Graduate Scholarship, the University of Guelph, and is currently supported by an Ontario HIV Treatment Network Junior Investigator Development Award.

Authors' Contributions

The Cruising Counts Research Team included the following members: Dr David Brennan, Dr Nathan Lachowsky, Gerardo Betancourt, Miss Vijaya Chikermane, Meredith Fraser, Duncan MacLachlan, Owen McEwen, James Murray, Daniel Pugh, Garfield Durrant, Richard Utama, and Dr Simon Rosser.

Conflicts of Interest

None declared.

References

1. Public Health Agency of Canada. 2012. Summary: Estimates of HIV prevalence and incidence in Canada, 2011 URL: http://www.phac-aspc.gc.ca/aids-sida/publication/survreport/estimat2011-eng.php [accessed 2015-02-09] [WebCite Cache ID 6WDqNPbe0]
2. Public Health Agency of Canada. 2013. At a glance - HIV and AIDS in Canada: Surveillance report to December 31st, 2012 URL: http://www.phac-aspc.gc.ca/aids-sida/publication/survreport/2012/dec/index-eng.php [accessed 2015-11-12] [WebCite Cache ID 6czDnvpEF]
3. Ontario HIV Treatment Network. Ontario Community HIV and AIDS Reporting Tool (OCHART): View From the Front Lines. 2012. URL: https://www.ochart.ca/documents/2013/VFTFL-2012-WEB.pdf [accessed 2015-02-09] [WebCite Cache ID 6WDqWOArD]
4. Rosser BR, Wilkerson JM, Smolenski DJ, Oakes JM, Konstan J, Horvath KJ, et al. The future of Internet-based HIV prevention: A report on key findings from the Men's INTernet (MINTS-I, II) Sex Studies. AIDS Behav 2011 Apr;15 Suppl 1:S91-S100 [FREE Full text] [doi: 10.1007/s10461-011-9910-5] [Medline: 21360127]
5. Gudelunas D. There's an App for that: The uses and gratifications of online social networks for gay men. Sex Cult 2012 Jan 14;16(4):347-365. [doi: 10.1007/s12119-012-9127-4]
6. Davis M, Hart G, Bolding G, Sherr L, Elford J. Sex and the Internet: Gay men, risk reduction and serostatus. Cult Health Sex 2006;8(2):161-174. [doi: 10.1080/13691050500526126] [Medline: 16641064]
7. Myers T, Allman D, Adam B, Alexander S, Blais M, Calzavara LM, et al. Male Call Canada Technical Report. 2013 Apr 13. URL: http://www.malecall.ca/MaleCall/wp-content/uploads/2012/03/Male-Call-Canada-Technical-Report_13April13_DRAFT.pdf [accessed 2015-02-09] [WebCite Cache ID 6WDrvLnWv]
8. Gilbert M, Hottes TS, Kerr T, Taylor D, Fairley CK, Lester R, et al. Factors associated with intention to use Internet-based testing for sexually transmitted infections among men who have sex with men. J Med Internet Res 2013;15(11):e254 [FREE Full text] [doi: 10.2196/jmir.2888] [Medline: 24240644]
9. Myers T, Remis R, Husbands W. Lambda Survey: M-Track Ontario Second Generation Surveillance. URL: http://www.actoronto.ca/setup.nsf/ActiveFiles/Lambda+Survey+%28Technical+Report%29/$file/lambda%20technical%20report%20FINAL.pdf [accessed 2015-02-09] [WebCite Cache ID 6WDsKsLwV]
10. Cooper A. Sexuality and the Internet: Surfing into the new millennium. Cyberpsychol Behav 1998 Jan;1(2):187-193. [doi: 10.1089/cpb.1998.1.187]
11. King SA. Internet gambling and pornography: Illustrative examples of the psychological consequences of communication anarchy. Cyberpsychol Behav 1999;2(3):175-193. [doi: 10.1089/cpb.1999.2.175] [Medline: 19178236]
12. Cooper A. Sex and the Internet: A Guide Book for Clinicians. New York, NY: Brunner-Routledge; 2002.
13. McKie RM, Lachowsky NJ, Milhausen RR. The positive impact of technology on young gay men's dating and sexual relationships in Canada: Results from a focus group study. J LGBT Youth 2015 Jan 09;12(1):19-38. [doi: 10.1080/19361653.2014.935552]
14. Landovitz RJ, Tseng CH, Weissman M, Haymer M, Mendenhall B, Rogers K. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. J Urban Health 2013 Aug;90(4):729-739. [doi: 10.1007/s11524-012-9766-7] [Medline: 22983721]
15. Ogilvie GS, Taylor DL, Trussler T, Marchand R, Gilbert M, Moniruzzaman A, et al. Seeking sexual partners on the Internet: A marker for risky sexual behaviour in men who have sex with men. Can J Public Health 2008;99(3):185-188. [Medline: 18615938]

16. Rhodes SD, Hergenrather KC, Duncan J, Vissman AT, Miller C, Wilkin AM, et al. A pilot intervention utilizing Internet chat rooms to prevent HIV risk behaviors among men who have sex with men. Public Health Rep 2010;125 Suppl 1:29-37 [FREE Full text] [Medline: 20408385]
17. Bolding G, Davis M, Sherr L, Hart G, Elford J. Use of gay Internet sites and views about online health promotion among men who have sex with men. AIDS Care 2004 Nov;16(8):993-1001. [doi: 10.1080/09540120412331292453] [Medline: 15511731]
18. Moskowitz DA, Melton D, Owczarzak J. PowerON: The use of instant message counseling and the Internet to facilitate HIV/STD education and prevention. Patient Educ Couns 2009 Oct;77(1):20-26. [doi: 10.1016/j.pec.2009.01.002] [Medline: 19217742]
19. Hooper S, Rosser BR, Horvath KJ, Oakes JM, Danilenko G, Men's INTernet Sex II (MINTS-II) Team. An online needs assessment of a virtual community: What men who use the Internet to seek sex with men want in Internet-based HIV prevention. AIDS Behav 2008 Nov;12(6):867-875 [FREE Full text] [doi: 10.1007/s10461-008-9373-5] [Medline: 18401701]
20. Mikolajczak M, Kok G, Hospers H. Queermasters: Developing a theory- and evidence-based Internet HIV-prevention intervention to promote HIV-testing among men who have sex with men (MSM). Appl Psychol 2008;57(4):681-697. [doi: 10.1111/j.1464-0597.2008.00342.x]
21. Sandstra IL, Gold F, Jones E, Harris P, Taylor D. Cyber outreach: ST/HIV education online. Can Nurse 2008 Jun;104(6):24-28, 30-31. [Medline: 18642498]
22. Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296 [FREE Full text] [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]
23. Rosser BR, Oakes JM, Konstan J, Hooper S, Horvath KJ, Danilenko GP, et al. Reducing HIV risk behavior of men who have sex with men through persuasive computing: Results of the Men's INTernet Study-II. AIDS 2010 Aug 24;24(13):2099-2107 [FREE Full text] [doi: 10.1097/QAD.0b013e32833c4ac7] [Medline: 20601853]
24. Carpenter KM, Stoner SA, Mikko AN, Dhanak LP, Parsons JT. Efficacy of a Web-based intervention to reduce sexual risk in men who have sex with men. AIDS Behav 2010 Jun;14(3):549-557 [FREE Full text] [doi: 10.1007/s10461-009-9578-2] [Medline: 19499321]
25. Bowen AM, Williams ML, Daniel CM, Clayton S. Internet based HIV prevention research targeting rural MSM: Feasibility, acceptability, and preliminary efficacy. J Behav Med 2008 Dec;31(6):463-477 [FREE Full text] [doi: 10.1007/s10865-008-9171-6] [Medline: 18770021]
26. Tooley L. CATIE. 2015. Cyberspace and cellphones: New frontiers for HIV prevention with gay, bisexual and other men who have sex with men URL: http://www.catie.ca/en/pif/fall-2015/cyberspace-and-cellphones-new-frontiers-hiv-prevention-gay-bisexual-and-other-men-who- [accessed 2015-09-13] [WebCite Cache ID 6bWKyLeGc]
27. Rosenberger JG, Reece M, Novak DS, Mayer KH. The Internet as a valuable tool for promoting a new framework for sexual health among gay men and other men who have sex with men. AIDS Behav 2011 Apr;15 Suppl 1:S88-S90. [doi: 10.1007/s10461-011-9897-y] [Medline: 21331800]
28. Hottes TS, Farrell J, Bondyra M, Haag D, Shoveller J, Gilbert M. Internet-based HIV and sexually transmitted infection testing in British Columbia, Canada: Opinions and expectations of prospective clients. J Med Internet Res 2012;14(2):e41 [FREE Full text] [doi: 10.2196/jmir.1948] [Medline: 22394997]
29. Landovitz RJ, Tseng C, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. J Urban Health 2013 Aug;90(4):729-739 [FREE Full text] [doi: 10.1007/s11524-012-9766-7] [Medline: 22983721]
30. Rietmeijer CA, McFarlane M. Web 2.0 and beyond: Risks for sexually transmitted infections and opportunities for prevention. Curr Opin Infect Dis 2009 Feb;22(1):67-71. [doi: 10.1097/QCO.0b013e328320a871] [Medline: 19532082]
31. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: A review. J Med Internet Res 2013;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]
32. Dohoo I, Martin W, Stryhn H. Methods in Epidemiologic Research. Charlottetown, PEI: VER Inc; 2012.
33. Grov C, Breslow AS, Newcomb ME, Rosenberger JG, Bauermeister JA. Gay and bisexual men's use of the Internet: Research from the 1990s through 2013. J Sex Res 2014;51(4):390-409 [FREE Full text] [doi: 10.1080/00224499.2013.871626] [Medline: 24754360]
34. Two-spirit people and sexual diversity. In: Finding Our Way: A Sexual and Reproductive Health Sourcebook for Aboriginal Communities. Ottawa, ON: Aboriginal Nurses Association of Canada and Planned Parenthood Federation of Canada; 2002:231-237.
35. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: A systematic review. J Med Internet Res 2014 May;16(5):e134 [FREE Full text] [doi: 10.2196/jmir.3393] [Medline: 24862459]

36. Li Y, Polk J, Plankey M. Online health-searching behavior among HIV-seropositive and HIV-seronegative men who have sex with men in the Baltimore and Washington, DC area. J Med Internet Res 2013 May;15(5):e78 [FREE Full text] [doi: 10.2196/jmir.2479] [Medline: 23644412]
37. Pedrana A, Hellard M, Gold J, Ata N, Chang S, Howard S, et al. Queer as F**k: Reaching and engaging gay men in sexual health promotion through social networking sites. J Med Internet Res 2013 Feb;15(2):e25 [FREE Full text] [doi: 10.2196/jmir.2334] [Medline: 23391459]

Abbreviations

AOR: adjusted odds ratio
ASO: AIDS service organization
CBO: community-based organization
EI: Employment Insurance
GB2M: gay, bisexual, and two-spirited men, as well as other men who have sex with men (sexual minority men)
LGBT: lesbian, gay, bisexual, and transgender
MISM: men who use the Internet to seek sex with other men
ODSP: Ontario Disability Support Program
OR: odds ratio
PEP: postexposure prophylaxis
PHA: people living with HIV/AIDS
STI: sexually transmitted infection
UAI: unprotected anal intercourse

Edited by G Eysenbach; submitted 03.04.15; peer-reviewed by R Schnall, T Falasinnu, S Chang; comments to author 09.07.15; revised version received 15.09.15; accepted 09.10.15; published 09.12.15.

Please cite as:
Brennan DJ, Lachowsky NJ, Georgievski G, Rosser BRS, MacLachlan D, Murray J, Cruising Counts Research Team
Online Outreach Services Among Men Who Use the Internet to Seek Sex With Other Men (MISM) in Ontario, Canada: An Online Survey
J Med Internet Res 2015;17(12):e277
URL: http://www.jmir.org/2015/12/e277/
doi:10.2196/jmir.4503
PMID:26681440

Original Paper

Consumer Use of “Dr Google”: A Survey on Health Information-Seeking Behaviors and Navigational Needs

Kenneth Lee[1], BPharm (Hons); Kreshnik Hoti[1], BPharm, PhD; Jeffery David Hughes[1], BPharm, GradDipPharm, MPharm, PhD; Lynne M Emmerton[1], BPharm (Hons), PhD

Curtin University, School of Pharmacy, Curtin University, Perth, Australia

Corresponding Author:
Lynne M Emmerton, BPharm (Hons), PhD
Curtin University
School of Pharmacy
Curtin University
GPO Box U1987
Perth, 6845
Australia
Phone: 61 892667352
Fax: 61 892662769
Email: lynne.emmerton@curtin.edu.au

Abstract

Background: The Internet provides a platform to access health information and support self-management by consumers with chronic health conditions. Despite recognized barriers to accessing Web-based health information, there is a lack of research quantitatively exploring whether consumers report difficulty finding desired health information on the Internet and whether these consumers would like assistance (ie, navigational needs). Understanding navigational needs can provide a basis for interventions guiding consumers to quality Web-based health resources.

Objective: We aimed to (1) estimate the proportion of consumers with navigational needs among seekers of Web-based health information with chronic health conditions, (2) describe Web-based health information-seeking behaviors, level of patient activation, and level of eHealth literacy among consumers with navigational needs, and (3) explore variables predicting navigational needs.

Methods: A questionnaire was developed based on findings from a qualitative study on Web-based health information-seeking behaviors and navigational needs. This questionnaire also incorporated the eHealth Literacy Scale (eHEALS; a measure of self-perceived eHealth literacy) and PAM-13 (a measure of patient activation). The target population was consumers of Web-based health information with chronic health conditions. We surveyed a sample of 400 Australian adults, with recruitment coordinated by Qualtrics. This sample size was required to estimate the proportion of consumers identified with navigational needs with a precision of 4.9% either side of the true population value, with 95% confidence. A subsample was invited to retake the survey after 2 weeks to assess the test-retest reliability of the eHEALS and PAM-13.

Results: Of 514 individuals who met our eligibility criteria, 400 (77.8%) completed the questionnaire and 43 participants completed the retest. Approximately half (51.3%; 95% CI 46.4-56.2) of the population was identified with navigational needs. Participants with navigational needs appeared to look for more types of health information on the Internet and from a greater variety of information sources compared to participants without navigational needs. However, participants with navigational needs were significantly less likely to have high levels of eHealth literacy (adjusted odds ratio=0.83, 95% CI 0.78-0.89, *P*<.001). Age was also a significant predictor (*P*=.02).

Conclusions: Approximately half of the population of consumers of Web-based health information with chronic health conditions would benefit from support in finding health information on the Internet. Despite the popularity of the Internet as a source of health information, further work is recommended to maximize its potential as a tool to assist self-management in consumers with chronic health conditions.

(J Med Internet Res 2015;17(12):e288) doi:10.2196/jmir.4345

KEYWORDS

online; health information; health literacy; patient activation; information seeking; information needs; Internet; chronic disease; patients; survey

Introduction

The Internet offers a wealth of information on numerous topics. Its pervasiveness in everyday life means it is a common source of information for many consumers [1]. Many consumers use it to obtain health-related information [2-5]. Accordingly, a number of studies have examined the role of the Internet in health care and its influence on the traditional relationship between consumers and their health professionals [6-10]. Traditionally, health professionals have been the primary source of health information, providing information through patient education [11]. Consumers are now afforded greater access to information, have greater potential to be more informed, and are able to play a greater role in caring for their health [11].

Consumers also play an important role in health care, particularly given a trend towards greater burden of chronic health conditions [12]. Such conditions often require daily self-management. In Australia, annual expenditure on chronic health conditions is estimated at AUD $11.0 billion [13]. Internationally, a number of chronic health conditions have been listed in the top 10 leading causes of mortality [14]. Consequently, initiatives should focus on supporting consumers with chronic health conditions to better manage their conditions.

The popularity of the Internet for health-related purposes enables its use to support self-management. Indeed, numerous studies have examined the popularity of Internet use as a source of health information [2,3,5,15-17]. In the United States, 80% of Internet users use it for health information [3]. It appears that the use of the Internet for health information is more popular in Internet users with chronic health conditions or disabilities compared to Internet users without chronic health conditions or disabilities [18]. While fewer data are available within the Australian context, a 2010 study [5] suggested that almost 80% of Internet users in Australia access the Internet for health information. Despite the popularity of its use for health information, a number of studies have identified barriers to accessing Web-based health information [19-22]. The volume of health information available on the Internet [19-22], the abundance of poor quality information [19,20,23], and the lack of strict publishing guidelines [19] are some examples. Furthermore, a 2001 review on consumers' Web-based health-information seeking identified factors contributing to potential misinformation and subsequent potential for harm if consumers were to access and act upon misleading information [19]. Hence, there is a need to better understand consumers' Web-based health information-seeking behaviors (HISB) to better support consumers in their self-management.

Numerous studies have explored the characteristics of consumers' Web-based HISB [3-5,23-44]. However, within the context of consumers with chronic health conditions, the majority of studies appear to focus on specific chronic health conditions [24,26,28-32,34-38], age [33,41], or ethnic groups [40], or they involve general populations that include consumers without chronic health conditions [3-5,23,25,27,39]. The applicability of findings from such studies to other populations may be limited. We believe that exploration of Web-based HISB in a population of health information consumers with a variety of chronic health conditions can facilitate identification of general characteristics or trends of Web-based HISB; such characteristics can then be compared to existing and future studies that focus on specific populations.

A qualitative study was recently conducted using consumers of Web-based health information who identified as having one or more chronic health conditions [20]. This study explored the Web-based HISB of its participants and identified a number of potentially related characteristics. However, the applicability of these characteristics to a wider population is unknown. While a large-scale quantitative study has explored the characteristics of consumers with chronic health conditions and the proportion of Internet and non-Internet users [42], no large-scale quantitative studies examine the breadth of HISB characteristics reported by consumers in the aforementioned qualitative study [20]. For example, previous studies have examined characteristics such as the frequency of Internet use for health-related purposes [45] and consumers' experiences with locating Web-based health information [46]. Within the context of health information consumers with a variety of chronic health conditions, characteristics of Web-based HISB such as the types of health information sought on the Internet and reasons for seeking Web-based health information have yet to be quantitatively determined.

Related to Web-based HISB, as identified by [20], are the concepts of health literacy, eHealth literacy, and patient activation. Numerous studies have identified health literacy [47-51] and eHealth literacy [52,53] as important skills in locating, accessing, and utilizing quality health information for health care management. Patient activation is defined as patients' belief that they "have important roles to play in self-managing care, collaborating with providers, and maintaining their health. They know how to manage their condition and maintain functioning and prevent health declines; and they have the skills and behavioral repertoire to manage their condition, collaborate with their health providers, maintain their health functioning, and access appropriate and high-quality care" [54]. Some evidence supports a statistically significant relationship between health literacy and patient activation [55,56]. However, to our knowledge, no study to date has examined the relationship between eHealth literacy and patient activation.

Despite the aforementioned barriers to acquiring desired Web-based health information, health information seeking using the Internet remains a prevalent activity. Thus, beyond understanding consumers' Web-based HISB, eHealth literacy, and patient activation, researchers have not yet explored whether consumers have difficulty finding, and indicate a desire for support to find, Web-based health information (ie, navigational needs). While findings from a qualitative study suggest a potential need for support interventions among consumers [20], the applicability of this finding to a wider population has yet to be determined. Furthermore, there are no studies examining potential determinants or predictors of navigational needs. Once an understanding of navigational needs and an estimate of the proportion of the population with navigational needs is ascertained, future studies can then explore consumers' preferences for support interventions within and between various populations, such as populations with specific chronic health conditions, which better support consumers in their self-management.

Thus, this study aims to address the following objectives: (1) estimate the proportion of consumers with navigational needs among consumers of Web-based health information living with chronic health conditions, (2) describe the following characteristics of consumers with navigational needs: Web-based HISB, patient activation, and eHealth literacy, and (3) explore variables predicting navigational needs of these consumers.

Methods

Overview

A Web-based questionnaire was developed via the Qualtrics platform to identify the proportion of consumers with navigational needs and to explore their demographics, Web-based HISB, eHealth literacy, and patient activation.

Ethical approval for this study was granted by the Curtin University Human Research Ethics Committee (HR06/2013).

Participants and Recruitment

The target population for this study was adult Web-based health information consumers with chronic health conditions residing in Australia. Participants were included in this study if they consented to the study and indicated they met the following criteria: (1) able to easily read and write in English, (2) aged 18 years or older, (3) use of the Internet to find information about their health, and (4) have at least one chronic health condition.

Recruitment was conducted by Qualtrics through their partnership with a Web-based survey research company, ResearchNow, which hosts a large diverse pool of participants and has the ability to select representative samples meeting specified eligibility criteria [57].

Sample Size

The sample size was determined using conservative parameters for prevalence studies [58]—our focus for prevalence estimation being the proportion of the target population with navigational needs (Objective 1). In the absence of literature reporting this prevalence, we used the following parameters: expected population proportion of 50%, 95% confidence interval, and a level of precision of estimate within 5% either side of the true population proportion. These parameters indicated a required sample size of 385 participants [58]. To account for potential invalid responses, the required sample was increased to 400 participants (a level of precision of 4.9% either side). This sample size was also deemed adequate to conduct descriptive and inferential statistical analyses to address the other objectives. The research company was contracted to meet the quota of 400 submitted questionnaires.

Questionnaire Development

Initial Questionnaire Construction

Questions and response items pertaining to navigational needs and Web-based HISB were predominantly drawn from interview questions and participant responses from a qualitative study [20] of health consumers with chronic health conditions who used the Internet. To ensure that questions asked verbally in the aforementioned qualitative study [20] were suitable for a written questionnaire, the wording of the interview questions was modified by the primary researcher with review from the other researchers on the research team. Similarly, decisions for choosing which interview questions were to be included as survey questions were made by the primary researcher in collaboration with the research team. Further items were added to supplement these questions and facilitate statistical analysis after discussion with all authors. Question types were a mix of 5-point Likert-type scales and multiple-response, multiple-choice options. Where relevant, multiple-choice items facilitated "other" responses to be typed and later manually coded for analysis. To mitigate the potential for selection bias within questions, the order of response items within each multiple-choice question was randomized where appropriate [59]. To reduce the number of questions and therefore respondent fatigue, adaptive questioning was used [59].

The eHealth Literacy Scale (eHEALS), a measure of perceived eHealth literacy [60], and PAM-13 [61], a measure of patient activation, were used to assess eHealth literacy and patient activation, respectively. Both of these scales had been assessed for validity and reliability [54,60-68] and were incorporated with permission from their respective authors/licensors.

Pilot Test

A target of 40 completed responses (10% of the final sample) was used to pilot test the questionnaire. Participants recruited for this stage were to meet the same eligibility criteria as our test sample and were recruited by Qualtrics via ResearchNow. Participants from the pilot sample were excluded from participation in the test sample to mitigate response bias.

The purposes of pilot testing were to assess comprehension of questions and response items and to examine questions with invalid or poor responses. Participants were encouraged, in space provided after each question, to provide comments regarding the comprehensibility of questions and response items.

Questionnaire Refinement

Based on participant feedback in the pilot test, a number of amendments were made to the wording of questions and response items, along with presentation of the questions for completion in Web-based format. First, the questionnaire enabled "attention-filter" questions; thus, response items were added to identify invalid responses (eg, "I am paying attention; please select 'disagree' for this line"). Three attention filters were included in this questionnaire: two questions instructed participants to select a certain option, and one response item instructed participants not to select the item. These attention filters were inserted into parts of the questionnaire that required longer attention spans (eg, long questions or questions with numerous response items). Second, wording of questions with lower response rates were revised, and these questions were marked as forced responses where possible to facilitate statistical analysis. To ensure participants were permitted to respond with “not applicable” for forced response questions, an “Other” option was provided wherever possible, with free-text space to explain their situation. Third, the mean survey completion time from the pilot test was relayed to Qualtrics to determine a “duration filter” for the test sample. The time parameter for the duration filter was calculated to be one-third of the mean pilot questionnaire completion time, as recommended by Qualtrics, and excluded participants who completed the questionnaire in a shorter-than-expected time. All questions and response items were examined by the research team to ensure readability and face validity prior to survey administration.

A Flesch-Kincaid Grade Level test [69] was conducted to test the readability of the questionnaire, including the informed consent and eligibility screening page, to compare to participants’ reported level of education.

Reliability Testing

A subset of 48 participants (approximately 10%, allowing extra in the case of delays in acceptance or questionnaire completion) was invited 2 weeks after completion of the questionnaire to retake the questionnaire, to confirm the test-retest reliability of the eHEALS and PAM-13 against reported values.

Analysis

Overview

All statistical analyses were conducted using SPSS version 21. Descriptive statistics were used to address Objective 1. Descriptive statistics, Pearson correlation, and multivariate linear regression were utilized to address Objective 2. Bivariate and multivariate binary logistic regressions were conducted to address Objective 3. Scores for the eHEALS and PAM-13 were calculated as per the authors’ instructions and were used in the regression modeling (Objectives 2 and 3).

All variables to be tested in the regression analyses were entered via a forced-entry method, as this method is more stable against random variation in the data, compared to other methods such as stepwise methods [70]. Demographic variables of age, sex, and level of education were entered alongside the other test variables, as these variables have been identified as potential contributors to the usage of Web-based health information [40,71]. The demographic variable examining residence in major cities or rural areas was also included in the regression model, as rurality has been identified as a potential barrier to Internet access [5]. Given the categorical nature of our demographic variables, categories with low or zero frequencies were aggregated with other categories, where logical, to allow for valid statistical conclusions. To illustrate, for the “age” variable (see Table 1), less than 1% of participants indicated that they were above the 55-64 years age category; the categories 65-74, 75-84, and 85+ were therefore combined with the 55-64 age category and relabeled as 55+ for inferential statistical analysis. Similarly, for the remoteness of residence variable (see Table 1), few participants indicated that they reside in remote areas; this category was aggregated with rural or regional areas to allow for comparison between major city areas versus rural/regional/remote. Such decisions for aggregating categories were made by the primary researcher in discussion with all other researchers within the research team. The level of significance (alpha) was set at $P<.05$.

Table 1. Demographic descriptors of respondents (N=400).

Category	No navigational needs (N=195), n (%)	Navigational needs (N=205), n (%)	Total, n (%)
Sex			
Male	73 (37.4)	82 (40.0)	155 (38.8)
Female	122 (62.6)	123 (60.0)	245 (61.3)
Age group (years)			
18-24	22 (11.3)	22 (10.7)	44 (11.0)
25-34	49 (25.1)	71 (34.6)	120 (30.0)
35-44	35 (17.9)	36 (17.6)	71 (17.8)
45-54	52 (26.7)	30 (14.6)	82 (20.5)
55-64	36 (18.5)	44 (21.5)	80 (20.0)
65-74	1 (0.5)	2 (1.0)	3 (0.8)
75-84	0 (0.0)	0 (0.0)	0 (0.0)
≥85	0 (0.0)	0 (0.0)	0 (0.0)
Level of formal education			
No formal education	0 (0.0)	0 (0.0)	0 (0.0)
Primary school	2 (1.0)	0 (0.0)	2 (0.5)
Junior high school	21 (10.8)	13 (6.3)	34 (8.5)
Senior high school	38 (19.5)	47 (22.9)	85 (21.3)
TAFE or technical college	53 (27.2)	62 (30.2)	115 (28.8)
University	81 (41.5)	83 (40.5)	164 (41.0)
Remoteness of residence			
Major city area	122 (62.6)	144 (70.2)	266 (66.5)
Rural or regional area	69 (35.4)	61 (29.8)	130 (32.5)
Remote area	4 (2.1)	0 (0.0)	4 (1.0)

Navigational Needs

The term “navigational needs” has been used above and refers to individuals who report having difficulty finding, and would like support in locating, desired Web-based health information. As no objective measure of navigational needs is available in the literature, we operationally defined the term as individuals who identified that they at least “sometimes” have difficulty locating desired Web-based health information (Criterion 1) and indicated that they would like help locating desired Web-based health information (Criterion 2).

These participants were considered a subset of the total respondents for the purposes of data analysis and were descriptively compared (Objectives 1 and 2). For Objective 3, this subset was compared to the remainder of the sample using binary logistic regression to determine predictors of navigational needs.

Reliability Tests

Statistical procedures to test the reliability of the eHEALS and PAM-13 [60,62,64-68] were replicated in our study (Objective 2, patient activation and eHealth literacy). These tests included internal consistency (Cronbach alpha) and intraclass correlation coefficient (ICC). ICC was assessed via a two-way mixed effects model [72] using an absolute agreement definition, ICC (3,1). This decision was made given the self-reported nature of our questionnaire and our intent to assess the agreement of participant responses to both PAM-13 and eHEALS between test and retest. Results from each of these tests were considered alongside relevant guidelines to assist interpretation [72,73].

Results

Summary

The survey was conducted during May 2014. In order to obtain our target of 400 submitted questionnaires, a total of 1104 individuals were invited by ResearchNow from their diverse participant pool. Of these 1104 individuals, 1027 agreed to participate (93.03% consent). Of the 1027 individuals, 514 individuals (50.05%) met our eligibility criteria, and 400 (77.82%) completed the questionnaire.

In the retest sample 2 weeks post-completion, 47 of the 48 participants contacted agreed to participate again (98% consent). Of these, 43 completed the questionnaire a second time (91% completion).

The Flesch-Kincaid Grade Level for our questionnaire, including the informed consent and eligibility screening questions, resulted in a readability score of 8.0.

Proportion of Consumers With Navigational Needs

As established above, participants were operationally defined as having navigational needs if they met both Criterions 1 and 2. To assess Criterion 1, participants were asked to rate, on a 5-point Likert-type scale (Never, Rarely, Sometimes, Most of the Time, Always), how often they have difficulty finding desired Web-based health information. A total of 216 participants (54.0%) indicated that they experienced difficulty at least sometimes, thereby meeting Criterion 1.

To assess Criterion 2, participants indicated whether they would like help finding desired Web-based health information. A total of 365 participants (91.3%) met this criterion.

A total of 205 participants (51.3%) met both Criteria 1 and 2 for navigational needs. The estimated proportion of consumers with navigational needs among consumers of Web-based health information living with chronic health conditions was thus estimated at 51.3% (95% CI 46.4%-56.2%).

Demographic Characteristics

Of the 400 participants, 61.3% were female (245/400), 41.0% reported having a university-level of education (164/400), and 66.5% (266/400) reported being located in major city areas within Australia (see Table 1). Descriptive comparisons of the demographics of participants with and without navigational needs are included in Table 1; significance testing of these comparisons as potential predictors of navigational needs is illustrated later. Overall, demographic characteristics between participants with and without navigational needs appear comparable (see Table 1). Noteworthy exceptions include a higher proportion of participants with navigational needs who were aged 25-34 years old compared to participants without navigational needs (34.6% vs 25.1%), and a lower proportion of participants with navigational needs who were aged 45-54 years old compared to participants without navigational needs (14.6% vs 26.7%). Reported chronic health conditions varied widely, with conditions involving the major organs most prevalent (see Multimedia Appendix 1).

Web-Based Health Information-Seeking Behaviors

Descriptive comparisons of the Web-based HISB between participants with and without navigational needs are provided in Tables 2-4 (as well as Multimedia Appendices 2-4). Significance testing was not performed, as multiple-response items did not allow variables to be analyzed independently.

The categories of health information reportedly sought varied considerably; however, participants with navigational needs appeared to look for more types of health information compared to participants without navigational needs (see Multimedia Appendix 2). Similarly, when comparing participants with and without navigational needs, participants with navigational needs appeared to use more sources of Web-based health information (see Multimedia Appendix 3).

Most commonly, participants sought information on the Internet to be more informed and engaged in their self-care (see Table 2). In comparing participants with and without navigational needs, we found that participants with navigational needs appear to seek Web-based health information because they are less satisfied with their health professionals, but less interested in wanting to manage their own conditions (see Table 2). However, more participants with navigational needs appeared to act on the acquired health information compared to participants without navigational needs (see Multimedia Appendix 4).

Table 2. Why Web-based health information is sought (N=400).

Reason for seeking Web-based health information	No navigational needs (N=195), n (%)[a]	Navigational needs (N=205), n (%)[a]	Total, n (%)[a]
I want to be more informed.	155 (79.5)	169 (82.4)	324 (81.0)
I want to help manage my own condition.	143 (73.3)	127 (62.0)	270 (67.5)
I want to clarify information that has been given to me by a health professional.	114 (58.5)	109 (53.2)	223 (55.8)
Just out of interest.	105 (53.8)	107 (52.2)	212 (53.0)
I want to check information that was discussed during a consultation with a health professional.	89 (45.6)	109 (53.2)	198 (49.5)
I want to look for alternative or additional treatment options.	94 (48.2)	98 (47.8)	192 (48.0)
I want to have information to read.	91 (46.7)	93 (45.4)	184 (46.0)
I find there is limited time during a consultation with a health professional.	48 (24.6)	69 (33.7)	117 (29.3)
I am not provided with enough information during a consultation with a health professional.	38 (19.5)	61 (29.8)	99 (24.8)
I disagree with certain points made by a health professional.	17 (8.7)	24 (11.7)	41 (10.3)
Other	12 (6.2)	11 (5.4)	23 (5.8)

[a] Respondents could select multiple options; percentages do not total 100%.

Most of the participants, 94.5% (378/400) reported that they discussed health information sourced on the Internet with health professionals at least some of the time. Reasons for this behavior are suggestive of seeking professional opinion, along with a desire to engage further in self-management (see Table 3). Such reasons for discussing Web-based health information with health professionals appear comparable between participants with and without navigational needs. A notable exception is that a greater proportion of participants with navigational needs have discussions with health professionals to "find out more information" compared to participants without navigational needs.

Table 3. Reasons why health information obtained on the Internet is discussed with health professionals (N=378).

Consultation with health professionals	No navigational needs (N=181), n (%)[a]	Navigational needs (N=197), n (%)[a]	Total, n (%)[a]
I want to get the health professional's opinion on information that I found on the Internet.	123 (68.0)	130 (66.0)	253 (66.9)
I want to find out more information.	98 (54.1)	121 (61.4)	219 (57.9)
I want to be in control of the management of my health condition(s).	95 (52.5)	101 (51.3)	196 (51.9)
I trust the health professional.	81 (44.8)	81 (41.1)	162 (42.9)
I want to discuss alternative treatments, tests, or procedures.	79 (43.6)	74 (37.6)	153 (40.5)
I want to clarify information that was unclear on the website(s) that I visited.	67 (37.0)	80 (40.6)	147 (38.9)
Other	3 (1.7)	5 (2.5)	8 (2.1)

[a] Respondents could select multiple options; percentages do not total 100%.

Similarly, 98.8% (395/400) of the participants reported that they do not discuss health information sourced from the Internet with health professionals at least some of the time. Common reasons reported for not always discussing Web-based health information with health professionals relate to not wanting to embarrass oneself in front of health professionals and a belief that health professionals do not have the time to discuss health information sought on the Internet (see Table 4).

Table 4. Reasons why health information obtained from the Internet may not be discussed with health professionals (N=395).

Reason for not discussing with health professionals	No navigational needs (N=193), n (%)[a]	Navigational needs (N=202), n (%)[a]	Total, n (%)[a]
I do not want to embarrass myself in front of my health professional.	58 (30.1)	75 (37.1)	133 (33.7)
I do not think that health professionals have enough time to discuss what I find on the Internet.	53 (27.5)	71 (35.1)	124 (31.4)
I feel that I have enough information already.	64 (33.2)	59 (29.2)	123 (31.1)
I do not want to upset my health professional.	35 (18.1)	52 (25.7)	87 (22.0)
Other	37 (19.2)	23 (11.4)	60 (15.2)

[a] Respondents could select multiple options; percentages do not total 100%.

Patient Activation and eHealth Literacy

Summary

Tables 5 and 6 describe the patient activation and eHealth literacy scores based on the PAM-13 and eHEALS, respectively.

Compared to participants without navigational needs, participants with navigational needs appear, on the whole, to be less activated (see Table 5) and have a lower level of eHealth literacy (see Table 6).

Table 5. Summary statistics: PAM-13 scores.

PAM-13 score (0.0-100.0)	No navigational needs (N=195)	Navigational needs (N=204)	Total (N=399)[a]
Mean (SD)	63.1 (12.5)	58.9 (13.3)	61.0 (13.1)
Median	60.6	58.1	58.1
Mode	55.6	63.1	55.6
Range	24.1-100.0	35.5-100.0	24.4-100.0

[a] Score could not be calculated for one participant due to invalid responses.

Table 6. Summary statistics: eHEALS scores.

eHEALS Score (8.0-40.0)	No navigational needs (N=195)	Navigational needs (N=205)	Total (N=400)
Mean (SD)	31.0 (4.1)	28.2 (4.2)	29.5 (4.3)
Median	31.0	28.0	30.0
Mode	32.0	32.0	32.0
Range	16.0-40.0	15.0-40.0	15.0-40.0

Associations

Correlations between PAM-13 and eHEALS scores revealed a positive, moderate association (*r*=.50, *P*<.001) (see Table 7). After inclusion of sex, age group (compared to the "55+" reference group), education (university vs no university level of education), and place of residence (major city vs rural) variables into a multivariate model, the only statistically significant predictor of PAM-13 scores was eHEALS scores (*P*<.001).

Table 7. PAM-13 score vs eHEALS score, and demographic variables.

	B	SE B	β
Constant	18.01	4.05	—
Age groups			
18-24 (vs 55+)	-1.48	2.19	-.04
25-34 (vs 55+)	-1.03	1.69	-.04
35-44 (vs 55+)	-2.89	1.89	-.08
45-54 (vs 55+)	-1.34	1.76	-.04
Female	-1.80	1.21	-.07
University education	1.45	1.21	.05
Living in major city	-0.85	1.27	-.03
eHEALS Score	1.53	0.13	0.51[a]
(R^2=.27, Adj. R^2=.25)			

[a] *P*<.001

Reliability Tests

Internal consistency for the PAM-13 and eHEALS were assessed via Cronbach alpha for the test sample (n=400) and the retest sample (n=43). Relative test-retest reliability was assessed using ICC (3,1) to assess the overall test-retest reliability of the subset of the test sample (ie, n=43 from n=400) on retest. Results for the reliability tests indicate good-to-excellent internal consistency and excellent test-retest reliability (see Table 8).

Table 8. Reliability statistics for the PAM-13 and eHEALS.

		Cronbach α	ICC (3,1) (95% CI)
eHEALS			
	Test (n=400)	.87	N/A
	Test/Retest (n=43)	.92/.91	.79 (0.65-0.88)
PAM-13			
	Test (n=400)	.86	N/A
	Test/Retest (n=43)	.92/.88	.86 (0.75-0.92)

Predictors of Navigational Needs

After inclusion of age, sex, level of education (university vs no university level education), place of residence (major city vs rural), the PAM-13 score, and the eHEALS score into a multivariate model, only age (*P*=.02)—specifically, the 45-54 age group (*P*=.048)—and the eHEALS score (*P*<.001) were statistically significant predictors of navigational needs (see Table 9).

Overall, the predictor variables (demographic variables, PAM-13, and eHEALS scores) used in this binary logistic regression analysis explained 18.7% of the variance in having navigational needs, measured using Nagelkerke's R^2[70].

Table 9. Predictors of navigational needs.

Predictors	Navigational needs (N=205)[a]	No navigational needs (N=195)[a]	OR (95% CI)	Adjusted OR (95% CI)
Age group				
18-24	22 (10.7)	22 (11.3)	0.80 (0.39-1.67)	0.94 (0.42-2.11)
25-34	71 (34.6)	48 (24.6)	1.17 (0.66-2.05)	1.54 (0.81-2.92)
35-44	36 (17.6)	35 (17.9)	0.83 (0.44-1.56)	0.96 (0.48-1.96)
45-54	30 (14.6)	52 (26.7)	0.46 (0.25-0.87)[b]	0.51 (0.26-0.99)[c]
55+ (reference group for “age group” variable)	46 (22.4)	37 (19.0)	—	—
Female	123 (60.0)	122 (62.6)	0.90 (0.60-1.34)	0.98 (0.62-1.55)
University education	83 (41.1)	81 (41.5)	0.96 (0.64-1.43)	0.95 (0.60-1.50)
Living in major city	145 (40.5)	121 (62.1)	1.41 (0.93-2.15)	1.33 (0.83-2.15)
eHEALS score, mean (SD)	28.2 (4.2)	31.0 (4.1)	0.84 (0.80-0.89)[d]	0.83 (0.78-0.89)[d]
PAM-13 score, mean (SD)	58.9 (13.3)	63.1 (12.5)	0.98 (0.96-0.99)[e]	1.00 (0.98-1.02)

[a] Values presented as n (%) unless otherwise noted.

[b] *P*=.02.

[c] *P*=.048.

[d] *P*<.001.

[e] *P*=.002.

Based on adjusted odds ratios (adjusted OR), participants aged 45-54 years old were 0.51 times as likely to have navigational needs compared to participants aged 55 years and above. In other words, participants aged 45-54 years old were less likely to have navigational needs compared to participants aged 55 years and above. In addition, participants with a lower eHEALS score, that is, lower eHealth literacy, were more likely to have navigational needs.

Discussion

Principal Findings

Approximately half the population (51.3%, 95% CI 46.4-56.2) of consumers seeking Web-based health information and living with chronic health conditions was estimated to have navigational needs. These consumers reported at least some difficulty locating desired health information and indicated preferences for guidance to find desired health information on the Internet. While age and perceived eHealth literacy levels were associated with consumers’ navigational needs (Table 9), our study suggests that the majority of the population (91.3%), including consumers who did not report having difficulty locating desired Web-based health information, would still like some form of guidance. Given that approximately 75% of the total population in Australia reported having at least one chronic health condition [74], close to 75% of the Australian population has Internet access [5], and nearly 80% of Internet users use the Internet for health-related activities [5], it appears that a sizeable proportion of the total Australian population would likely be willing to receive some form of guidance in locating desired Web-based health information. A previous qualitative study [20] suggested health professionals could play a role in helping consumers locate desired Web-based health information. This and other types of assistance will be explored elsewhere using our current data.

Findings from this study suggest consumers with at least one chronic condition want to be more informed about their health, and consumers seek information as a way to help manage their conditions. These findings support literature on the use of the Internet as a mechanism by health consumers to assist self-management [6,7,11,26]. When comparing participants with and without navigational needs, this study found that participants with navigational needs appear to look for more types of Web-based health information and from a greater variety of sources. Thus, this study adds to existing literature by providing some descriptive characteristics about the Web-based HISB of consumers with navigational needs.

The majority of participants reported discussing Web-based health information with health professionals; the most common reason was to ascertain the opinions of health professionals on the retrieved health information. Only 10.3% of our participants indicated they use the Internet to find health information when disagreeing with advice from their health professionals. While not underestimating the proportion of these participants, this suggests consumers living with chronic health conditions predominantly use Web-based health information for reasons other than overriding advice given by health professionals. Collectively, these findings appear in line with studies [6-8] examining the role of the Internet in the consumer-health professional relationship, in that the Internet has the potential to better facilitate this relationship. However, when comparing participants with and without navigational needs, this study found that participants with navigational needs were less likely to be satisfied with their health professionals and more likely to not discuss information with their health professionals because they did not want to embarrass themselves in front of their health professionals (see Table 4). Thus, this study provides initial insight into aspects of HISB in consumers with navigational needs and suggests that health professionals may need to have conversations with consumers that greater encourage discussion of health information sought using the Internet.

In our study, when compared to participants without navigational needs, participants with navigational needs appeared to have lower levels of patient activation and eHealth literacy (Tables 5 and 6). However, as established earlier, participants with navigational needs sought more types of health information from a greater variety of Web-based sources. Furthermore, participants with navigational needs were more likely to report that they discussed information sought using the Internet with their health professionals for the purpose of obtaining more information (Table 3). Such findings suggest that, despite have a seemingly greater desire to obtain information, participants with navigational needs are less able to find such information (lower eHealth literacy) and are less confident in their searching abilities (lower patient activation). This reinforces the need to provide assistance to consumers with navigational needs and provides further justification that more research needs to be conducted to address navigational needs.

Further to our use of the PAM-13 and eHEALS measures, our data revealed a moderate but statistically significant correlation between the PAM-13 and eHEALS (r=.50, P<.001), supporting a relationship between patient activation and perceived eHealth literacy, as well as confirming other studies [55,56]. These findings extend the literature in that patient activation appears to be a prominent concept in the context of eHealth literacy and suggests this association is present even after accounting for demographic variables of age, sex, level of education, and place of residence (major cities versus rural). While self-perceived eHealth literacy refers to individuals' self-perceived abilities to obtain and utilize Web-based health information for the purpose of self-management [52], patient activation refers to individuals' self-belief in their behavioral repertoires, abilities, and knowledge pertaining to self-management [54]. Given the apparent overlap in these two concepts, health information consumers who self-identify as being motivated and having the ability and knowledge to self-manage their conditions could also be assumed to be more adept at utilizing the Internet for self-management purposes.

The validity and reliability of the eHEALS and PAM-13 have been well established [54,60-68]. Our internal consistency and test-retest reliability analysis confirmed the reliability of both instruments in the current sample. Given that reliability is a prerequisite for validity [75], and the pre-establishment of validity in these two measures, these measures are likely to also be valid in our sample. By using these two measures as proxies for key concepts in predicting navigational needs, we believe our conclusions regarding the predictors of navigational needs are empirically justified.

Strengths and Limitations

A key strength of this study lies in our overall approach to developing the questionnaire. Specifically, the use of attention filters and a duration filter helped ensure that our participants provided complete and valid responses. The incorporation of two scales (PAM-13 and eHEALS) with prior evidence of validity and reliability allowed for trustworthy conclusions to be drawn from the data. The use of forced responses minimized potential for missing data; only one participant's PAM-13 score could not be calculated from having selected several "Not Applicable" options within the scale. The use of questions and response items pertaining to Web-based HISB and navigational needs, based on a qualitative study conducted on a similar target population [20], provided initial empirical validation. This means the characteristics of consumers' Web-based HISB and navigational needs from our study may more accurately reflect the target population, compared to questionnaires where the items had not been created from the consumer perspective. Furthermore, our use of randomization of response items perceivably mitigated response bias.

Various steps ensured our questions were easily comprehended by our participants: inviting participants to comment on comprehension in our pilot survey, face validity checks by our research team, and the use of the Flesch-Kincaid Grade level test. The readability score of 8.0 indicates participants who have completed at least the 8th grade of formal education would be able to comprehend the questionnaire [69]. Based on the demographics of our participants, this suggests that 99.5% of our participants (398/400 participants) would have been able to comprehend the questions and response items in our questionnaire.

While we requested Qualtrics to gather a representative sample of the Australian population, information pertaining to their sampling technique in doing so was not disclosed. As established earlier, there is a lack of data on population demographics in the context of Australian Internet use, limiting our ability to compare this sample with national demographic data. Furthermore, a representative sample does not necessarily translate to a random sample. Given our sample size calculation for prevalence studies assumes random sampling technique, the level of precision for our study cannot be accurately determined, and external validity cannot be assured. However, based on a report [76] from the Australian Institute of Health and Welfare (AIHW), the population prevalence versus the prevalence in our sample is comparable for cardiovascular diseases (22% vs 29.8%) and mental health conditions (20% vs 25.5%)—two of the three most commonly reported conditions in our sample (Table 2). This comparison, however, does not take into consideration variation in the prevalence of such conditions between Internet users and non-users. Furthermore, our sample is of consumers with chronic health conditions, whereas the AIHW report [76] expresses prevalence as a proportion of the entire Australian population. Nevertheless, our sample size was sufficient for the required analyses, and given the moderately large sample size and the diverse demographic characteristics of our participants, it appears our findings can be applied to a wider population.

A further limitation to our study was the use of multiple-response, multiple-choice questions for our Web-based HISB domain to generate a comprehensive description of this domain. The permutations of options meant the data were only reported descriptively; a much larger sample would be required to facilitate comparisons between cohorts of respondents. Similarly, test-retest reliability of such multiple-response, multiple-choice questions could not be determined.

This study did not determine the device(s) or platform(s) used to access Web-based health information. The availability of mobile-friendly versions of certain websites should improve access to Web-based health information [77]. In addition, our questionnaire did not explicitly explore social media as a health information source. These responses were elicited via an “Other” option (Table 5), but not to the extent suggested in literature reporting social network sites are becoming popular sources of general information for many users [42]. While such social network sites are reportedly less popular for people with chronic health conditions in the United States [42], their use is reportedly increasing [78].

Further Research

Despite the perception of the health professional as the most trusted source of Web-based health information [79], our study suggests that their role in guiding consumers to Web-based health information appears underutilized by consumers. We therefore recommend further investigation into why this role is underutilized and believe that both consumer and health professional perspectives should be explored. Once more in-depth understanding is attained, further research could explore the current roles of various health professions and investigate pragmatic ways that navigational guidance can be provided. While initiatives such as the use of social networking technologies by health professionals to provide guidance [80] and “information prescriptions” [81-85] have been implemented, to our knowledge, these initiatives have not considered consumers’ navigational needs, and this represents a topic for future development.

While age and perceived eHealth literacy levels were found to be statistically significant predictors of navigational needs, the variables included in our multivariate binary logistic regression model explain only 18.7% of the variance in a consumer being identified with navigational needs. Other variables associated with the navigational needs of consumers remain unexplored, and these may inform individualized approaches to supporting navigational needs. We therefore recommend further investigation into identifying additional predictors of navigational needs.

One could expect consumers with higher levels of perceived eHealth literacy would be less likely to have navigational needs. Indeed, this was the case in our study. Our study also identified a significantly lower likelihood of navigational needs in participants aged 45-54 years, compared to those aged 55 years and above. Further investigation is recommended to determine characteristics of this middle-aged group, why this specific age group was less likely to have navigational needs compared to those 55 years and above, and suitable interventions to meet their needs.

A positive correlation was found between patient activation and eHealth literacy, albeit moderate. Until empirical data can better account for the variance in the relationship, future interventions aimed to address either patient activation or eHealth literacy should retain both constructs. Finally, given our universal approach to exploring Web-based HISB of health information consumers with a variety of chronic health conditions, future studies that focus on specific chronic health conditions can compare their findings against this study to determine commonalities and variations between and across chronic conditions.

Conclusions

This study highlights the proportion of people with chronic health conditions who use the Internet and who have navigational needs, and reports that a majority of this population would want help locating desired Web-based health information. While we identified a number of associations that help identify individuals who would benefit from guidance in navigating Web-based health information, given that the majority of the population would want assistance, more universal approaches may be valuable to help all consumers locate desired Web-based health information.

Acknowledgments

The authors acknowledge the licensor of the PAM-13 (Insignia Health) and the authors of the eHEALS for permission to use both scales in our questionnaire. We also acknowledge Qualtrics staff for coordinating recruitment and providing technical support, and Dr Richard Parsons for statistical advice. KL is supported by an Australian Postgraduate Award.

Authors' Contributions

KL conceived and designed the study and the questionnaire with assistance from LE, KH, and JH. KL liaised with Qualtrics to coordinate recruitment. KL was involved with statistical analyses, interpretating findings, and drafting the manuscript. All authors were involved in reviewing and revising the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Reported chronic health conditions (N=400).

[PDF File (Adobe PDF File), 28KB - jmir_v17i12e288_app1.pdf]

Multimedia Appendix 2

Health information sought (N=400).

[PDF File (Adobe PDF File), 26KB - jmir_v17i12e288_app2.pdf]

Multimedia Appendix 3

Where Web-based health information is usually sought (N=400).

[PDF File (Adobe PDF File), 29KB - jmir_v17i12e288_app3.pdf]

Multimedia Appendix 4

Action(s) taken upon finding Web-based health information (N=400).

[PDF File (Adobe PDF File), 25KB - jmir_v17i12e288_app4.pdf]

References

1. Internet World Stats. 2014. Internet Usage Statistics: The Internet Big Picture - World Internet Users and 2014 Population Stats URL: http://www.internetworldstats.com/stats.htm [accessed 2014-01-13] [WebCite Cache ID 6WHVX82M1]
2. Fox S. Pew Internet & American Life Project. 2006. Online Health Search 2006 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2006/PIP_Online_Health_2006.pdf.pdf [accessed 2014-07-12] [WebCite Cache ID 6R0OwsHEC]
3. Fox S. Pew Internet & American Life Project. 2011. Health Topics URL: http://pewinternet.org/Reports/2011/HealthTopics.aspx [accessed 2012-08-07] [WebCite Cache ID 6SnedkDik]
4. Fox S, Jones S. Pew Internet & American Life Project. 2009. The Social Life of Health Information URL: http://www.pewinternet.org/2009/06/11/the-social-life-of-health-information/ [accessed 2014-07-12] [WebCite Cache ID 6R0Q6Jqpa]
5. McDaid D, Park A. Bupa. 2010. Online Health: Untangling the Web URL: https://www.bupa.com.au/staticfiles/Bupa/HealthAndWellness/MediaFiles/PDF/LSE_Report_Online_Health.pdf [accessed 2014-09-23] [WebCite Cache ID 6Snf9eDha]
6. Iverson SA, Howard KB, Penney BK. Impact of internet use on health-related behaviors and the patient-physician relationship: a survey-based study and review. J Am Osteopath Assoc 2008 Dec;108(12):699-711. [Medline: 19075034]
7. McMullan M. Patients using the Internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
8. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]
9. Hart A, Henwood F, Wyatt S. The role of the Internet in patient-practitioner relationships: findings from a qualitative research study. J Med Internet Res 2004 Sep 30;6(3):e36 [FREE Full text] [doi: 10.2196/jmir.6.3.e36] [Medline: 15471762]
10. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients’ use of the internet for medical information. J Gen Intern Med 2002 Mar;17(3):180-185. [doi: 10.1046/j.1525-1497.2002.10603.x]
11. Hoving C, Visser A, Mullen PD, van den Borne B. A history of patient education by health professionals in Europe and North America: from authority to shared decision making education. Patient Educ Couns 2010 Mar;78(3):275-281. [doi: 10.1016/j.pec.2010.01.015] [Medline: 20189746]
12. Donald M, Ware RS, Ozolins IZ, Begum N, Crowther R, Bain C. The role of patient activation in frequent attendance at primary care: a population-based study of people with chronic disease. Patient Educ Couns 2011 May;83(2):217-221. [doi: 10.1016/j.pec.2010.05.031] [Medline: 20598825]
13. AIHW. Australian Institute of Health and Welfare. Canberra; 2004. Health system expenditure on disease and injury in Australia, 2000-01 URL: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=6442457385 [accessed 2015-02-12] [WebCite Cache ID 6WHWJilGU]
14. World Health Organization. 2012. The top 10 causes of death URL: http://www.who.int/mediacentre/factsheets/fs310/en/ [accessed 2015-02-12] [WebCite Cache ID 6WHWOIT9U]
15. Andreassen HK, Bujnowska-Fedak MM, Chronaki CE, Dumitru RC, Pudule I, Santana S, et al. European citizens' use of E-health services: a study of seven countries. BMC Public Health 2007;7:53 [FREE Full text] [doi: 10.1186/1471-2458-7-53] [Medline: 17425798]
16. Fox S. Pew Internet & American Life Project. 2011. The Social Life of Health Information, 2011 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2011/PIP_Social_Life_of_Health_Info.pdf [accessed 2015-08-31] [WebCite Cache ID 6bBwjObNp]
17. Fox S. Pew Internet & American Life Project. 2011. Peer-to-Peer Healthcare URL: http://pewinternet.org/Reports/2011/P2PHealthcare.aspx [accessed 2012-08-07] [WebCite Cache ID 6bBwsz5UU]

18. Fox S. Pew Internet & American Life Project. 2007. E-Patients With a Disability or Chronic Disease URL: http://www.pewinternet.org/2007/10/08/e-patients-with-a-disability-or-chronic-disease/ [accessed 2015-08-31] [WebCite Cache ID 6bBwz7VEh]
19. Cline RJ, Haynes KM. Consumer health information seeking on the Internet: the state of the art. Health Educ Res 2001 Dec;16(6):671-692 [FREE Full text] [Medline: 11780707]
20. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]
21. McGrath I. Information superhighway or information traffic jam for healthcare consumers? Clin Perform Qual Health Care 1997;5(2):90-93. [Medline: 10167219]
22. Metzger MJ, Flanagin AJ. Using Web 2.0 technologies to enhance evidence-based medical information. J Health Commun 2011;16 Suppl 1:45-58. [doi: 10.1080/10810730.2011.589881] [Medline: 21843095]
23. Anderson JG. Consumers of e-Health: Patterns of Use and Barriers. Soc Sci Comput Rev 2004 May 01;22(2):242-248. [doi: 10.1177/0894439303262671]
24. Burkell JA, Wolfe DL, Potter PJ, Jutai JW. Information needs and information sources of individuals living with spinal cord injury. Health Info Libr J 2006 Dec;23(4):257-265 [FREE Full text] [doi: 10.1111/j.1471-1842.2006.00686.x] [Medline: 17177946]
25. Caiata-Zufferey M, Abraham A, Sommerhalder K, Schulz PJ. Online health information seeking in the context of the medical consultation in Switzerland. Qual Health Res 2010 Aug;20(8):1050-1061. [doi: 10.1177/1049732310368404] [Medline: 20442347]
26. Chou WS, Liu B, Post S, Hesse B. Health-related Internet use among cancer survivors: data from the Health Information National Trends Survey, 2003-2008. J Cancer Surviv 2011 Sep;5(3):263-270. [doi: 10.1007/s11764-011-0179-5] [Medline: 21505861]
27. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
28. James N, Daniels H, Rahman R, McConkey C, Derry J, Young A. A study of information seeking by cancer patients and their carers. Clin Oncol (R Coll Radiol) 2007 Jun;19(5):356-362. [doi: 10.1016/j.clon.2007.02.005] [Medline: 17399963]
29. Kalichman SC, Benotsch EG, Weinhardt L, Austin J, Luke W, Cherry C. Health-related Internet use, coping, social support, and health indicators in people living with HIV/AIDS: preliminary results from a community survey. Health Psychol 2003 Jan;22(1):111-116. [Medline: 12558209]
30. Lee SY, Hawkins R. Why do patients seek an alternative channel? The effects of unmet needs on patients' health-related Internet use. J Health Commun 2010 Mar;15(2):152-166. [doi: 10.1080/10810730903528033] [Medline: 20390984]
31. Matter B, Feinberg M, Schomer K, Harniss M, Brown P, Johnson K. Information needs of people with spinal cord injuries. J Spinal Cord Med 2009;32(5):545-554 [FREE Full text] [Medline: 20025150]
32. McCaughan E, McKenna H. Information-seeking behaviour of men newly diagnosed with cancer: a qualitative study. J Clin Nurs 2007 Nov;16(11):2105-2113. [doi: 10.1111/j.1365-2702.2006.01785.x] [Medline: 17331096]
33. Soederberg Miller LM, Bell RA. Online health information seeking: the influence of age, information trustworthiness, and search challenges. J Aging Health 2012 Apr;24(3):525-541. [doi: 10.1177/0898264311428167] [Medline: 22187092]
34. Noh H, Lee JM, Yun YH, Park SY, Bae D, Nam J, et al. Cervical cancer patient information-seeking behaviors, information needs, and information sources in South Korea. Support Care Cancer 2009 Oct;17(10):1277-1283. [doi: 10.1007/s00520-009-0581-y] [Medline: 19172301]
35. O'Grady L. Meeting health information needs of people with HIV/AIDS: sources and means of collaboration. Health Info Libr J 2008 Dec;25(4):261-269 [FREE Full text] [doi: 10.1111/j.1471-1842.2007.00764.x] [Medline: 19076672]
36. Roberts KJ, Lepore SJ, Urken ML. Quality of life after thyroid cancer: an assessment of patient needs and preferences for information and support. J Cancer Educ 2008;23(3):186-191. [doi: 10.1080/08858190802247762] [Medline: 18709591]
37. Samal L, Saha S, Chander G, Korthuis PT, Sharma RK, Sharp V, et al. Internet health information seeking behavior and antiretroviral adherence in persons living with HIV/AIDS. AIDS Patient Care STDS 2011 Jul;25(7):445-449 [FREE Full text] [doi: 10.1089/apc.2011.0027] [Medline: 21682586]
38. Tustin N. The role of patient satisfaction in online health information seeking. J Health Commun 2010 Jan;15(1):3-17. [doi: 10.1080/10810730903465491] [Medline: 20390974]
39. Yan YY. Online health information seeking behavior in Hong Kong: an exploratory study. J Med Syst 2010 Apr;34(2):147-153. [Medline: 20433053]
40. Lee YJ, Boden-Albala B, Larson E, Wilcox A, Bakken S. Online health information seeking behaviors of Hispanics in New York City: a community-based cross-sectional study. J Med Internet Res 2014;16(7):e176 [FREE Full text] [doi: 10.2196/jmir.3499] [Medline: 25092120]
41. Medlock S, Eslami S, Askari M, Arts DL, Sent D, de Rooij SE, et al. Health information-seeking behavior of seniors who use the Internet: a survey. J Med Internet Res 2015;17(1):e10 [FREE Full text] [doi: 10.2196/jmir.3749] [Medline: 25574815]
42. Fox S, Purcell K. Pew Internet & American Life Project. 2010. Chronic Disease and the Internet URL: http://pewinternet.org/Reports/2010/Chronic-Disease.aspx [accessed 2012-08-07] [WebCite Cache ID 6WHWsjzEf]

43. Rowlands IJ, Loxton D, Dobson A, Mishra GD. Seeking Health Information Online: Association with Young Australian Women's Physical, Mental, and Reproductive Health. J Med Internet Res 2015;17(5):e120 [FREE Full text] [doi: 10.2196/jmir.4048] [Medline: 25986630]
44. Jamal A, Khan SA, AlHumud A, Al-Duhyyim A, Alrashed M, Bin SF, et al. Association of Online Health Information-Seeking Behavior and Self-Care Activities Among Type 2 Diabetic Patients in Saudi Arabia. J Med Internet Res 2015;17(8):e196 [FREE Full text] [doi: 10.2196/jmir.4312] [Medline: 26268425]
45. Ayers SL, Kronenfeld JJ. Chronic illness and health-seeking information on the Internet. Health (London) 2007 Jul;11(3):327-347. [doi: 10.1177/1363459307077547] [Medline: 17606698]
46. Mayoh J, Bond CS, Todres L. An Innovative Mixed Methods Approach to Studying the Online Health Information Seeking Experiences of Adults With Chronic Health Conditions. Journal of Mixed Methods Research 2011 Aug 26;6(1):21-33. [doi: 10.1177/1558689811416942]
47. Baker DW, Wolf MS, Feinglass J, Thompson JA, Gazmararian JA, Huang J. Health literacy and mortality among elderly persons. Arch Intern Med 2007 Jul 23;167(14):1503-1509. [doi: 10.1001/archinte.167.14.1503] [Medline: 17646604]
48. Cutilli CC, Bennett IM. Understanding the health literacy of America: results of the National Assessment of Adult Literacy. Orthop Nurs 2009;28(1):27-32; quiz 33 [FREE Full text] [doi: 10.1097/01.NOR.0000345852.22122.d6] [Medline: 19190475]
49. Manning DL, Dickens C. Health literacy: more choice, but do cancer patients have the skills to decide? Eur J Cancer Care (Engl) 2006 Dec;15(5):448-452. [doi: 10.1111/j.1365-2354.2006.00687.x] [Medline: 17177901]
50. Nielsen-Bohlman L, Panzer AM, Kindig DA. Health literacy: a prescription to end confusion. Washington, DC: National Academies Press; 2004.
51. World Health Organization. 2015. Health literacy and health behavior URL: http://www.who.int/healthpromotion/conferences/7gchp/track2/en/ [accessed 2015-01-02] [WebCite Cache ID 6WHWyYx3c]
52. Norman CD, Skinner HA. eHealth Literacy: Essential Skills for Consumer Health in a Networked World. J Med Internet Res 2006;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
53. Collins SA, Currie LM, Bakken S, Vawdrey DK, Stone PW. Health literacy screening instruments for eHealth applications: a systematic review. J Biomed Inform 2012 Jun;45(3):598-607 [FREE Full text] [doi: 10.1016/j.jbi.2012.04.001] [Medline: 22521719]
54. Hibbard JH, Stockard J, Mahoney ER, Tusler M. Development of the Patient Activation Measure (PAM): conceptualizing and measuring activation in patients and consumers. Health Serv Res 2004 Aug;39(4 Pt 1):1005-1026 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00269.x] [Medline: 15230939]
55. Greene J, Hibbard J, Tusler M. How much do health literacy and patient activation contribute to older adults' ability to manage their health?. Washington, DC: AARP Public Policy Institute; 2005. URL: https://assets.aarp.org/rgcenter/health/2005_05_literacy.pdf [accessed 2015-02-12] [WebCite Cache ID 6WHX7laTb]
56. Smith SG, Curtis LM, Wardle J, von Wagner C, Wolf MS. Skill set or mind set? Associations between health literacy, patient activation and health. PLoS One 2013;8(9):e74373 [FREE Full text] [doi: 10.1371/journal.pone.0074373] [Medline: 24023942]
57. Research Now. 2015. Panels URL: http://www.researchnow.com/en-US/Panels.aspx [accessed 2015-02-12] [WebCite Cache ID 6WHXKar4X]
58. Arya R, Antonisamy B, Kumar S. Sample size estimation in prevalence studies. Indian J Pediatr 2012 Nov;79(11):1482-1488. [doi: 10.1007/s12098-012-0763-3] [Medline: 22562262]
59. 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] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
60. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
61. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. Health Serv Res 2005 Dec;40(6 Pt 1):1918-1930 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00438.x] [Medline: 16336556]
62. Brenk-Franz K, Hibbard JH, Herrmann WJ, Freund T, Szecsenyi J, Djalali S, et al. Validation of the German version of the patient activation measure 13 (PAM13-D) in an international multicentre study of primary care patients. PLoS One 2013;8(9):e74786 [FREE Full text] [doi: 10.1371/journal.pone.0074786] [Medline: 24098669]
63. Fowles JB, Terry P, Xi M, Hibbard J, Bloom CT, Harvey L. Measuring self-management of patients' and employees' health: further validation of the Patient Activation Measure (PAM) based on its relation to employee characteristics. Patient Educ Couns 2009 Oct;77(1):116-122. [doi: 10.1016/j.pec.2009.02.018] [Medline: 19356881]
64. Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form Patient Activation Measure (PAM13) in a Danish version. BMC Public Health 2009;9:209 [FREE Full text] [doi: 10.1186/1471-2458-9-209] [Medline: 19563630]
65. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in The Netherlands: translation and validation of the American short form Patient Activation Measure (PAM13). BMC Public Health 2012;12:577 [FREE Full text] [doi: 10.1186/1471-2458-12-577] [Medline: 22849664]

66. Stepleman L, Rutter M, Hibbard J, Johns L, Wright D, Hughes M. Validation of the patient activation measure in a multiple sclerosis clinic sample and implications for care. Disabil Rehabil 2010;32(19):1558-1567. [doi: 10.3109/09638280903567885] [Medline: 20590506]
67. Koo M, Norman CD, Chang HM. Psychometric evaluation of a Chinese version of the eHealth Literacy Scale (eHEALS) in school age children. Int Electron J Health Educ 2012 Dec 27;15:29-36 [FREE Full text]
68. Magnezi R, Glasser S. Psychometric properties of the hebrew translation of the patient activation measure (PAM-13). PLoS One 2014;9(11):e113391 [FREE Full text] [doi: 10.1371/journal.pone.0113391] [Medline: 25411841]
69. Kincaid J, Fishburne RP, Rogers RL, Chissom BS. National Technical Information Service. Springfield; 1975. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel URL: http://www.dtic.mil/dtic/tr/fulltext/u2/a006655.pdf [accessed 2015-02-12] [WebCite Cache ID 6WHY4cD2S]
70. Field A. Discovering statistics using SPSS. Los Angeles, CA: SAGE; 2009.
71. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of eHealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
72. Portney LG, Watkins MP. Foundations of clinical research: applications to practice. Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
73. Fleiss JL. The design and analysis of clinical experiments. New York: Wiley; 1999.
74. Australian Bureau of Statistics. 2010. National Health Survey: Summary of Results, 2007-08 URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/mf/4364.0 [accessed 2012-08-07] [WebCite Cache ID 6bC7QhnZ5]
75. Mitchell M, Jolley JM. Research Design Explained. Boston, MA: Wadsworth Publishing; 2013.
76. Australian Institute of Health and Welfare. Canberra; 2015. Chronic Diseases URL: http://www.aihw.gov.au/chronic-diseases/ [accessed 2015-02-12] [WebCite Cache ID 6bW4eOIAg]
77. Shrestha S. Mobile web browsing: usability study. In: MC07. 2007 Presented at: 4th Mobility Conference: International Conference on Mobile Technology, Applications, and Systems; Sept 12-14, 2007; Singapore. [doi: 10.1145/1378063.1378094]
78. Sadah SA, Shahbazi M, Wiley MT, Hristidis V. A Study of the Demographics of Web-Based Health-Related Social Media Users. J Med Internet Res 2015;17(8):e194 [FREE Full text] [doi: 10.2196/jmir.4308] [Medline: 26250986]
79. Dutta-Bergman M. Trusted online sources of health information: differences in demographics, health beliefs, and health-information orientation. J Med Internet Res 2003;5(3):e21 [FREE Full text] [doi: 10.2196/jmir.5.3.e21] [Medline: 14517112]
80. Young SD. Recommendations for Using Online Social Networking Technologies to Reduce Inaccurate Online Health Information. Online J Health Allied Sci 2011 Jul 30;10(2):- [FREE Full text] [Medline: 22016578]
81. Hand A, Greenwell K, Corbett S, Walker R, Kirrane H. Developing an information prescription service. Nurs Times 2008;104(39):34-37. [Medline: 19051547]
82. McKnight M. Information prescriptions, 1930-2013: an international history and comprehensive review. J Med Libr Assoc 2014 Oct;102(4):271-280 [FREE Full text] [doi: 10.3163/1536-5050.102.4.008] [Medline: 25349545]
83. Oliver KB, Lehmann HP, Wolff AC, Davidson LW, Donohue PK, Gilmore MM, et al. Evaluating information prescriptions in two clinical environments. J Med Libr Assoc 2011 Jul;99(3):237-246 [FREE Full text] [doi: 10.3163/1536-5050.99.3.011] [Medline: 21753916]
84. Ritterband LM, Borowitz S, Cox DJ, Kovatchev B, Walker LS, Lucas V, et al. Using the internet to provide information prescriptions. Pediatrics 2005 Nov;116(5):e643-e647. [doi: 10.1542/peds.2005-0404] [Medline: 16263978]
85. Siegel ER, Logan RA, Harnsberger RL, Cravedi K, Krause JA, Lyon B, et al. Information Rx: Evaluation of a new informatics tool for physicians, patients, and libraries. Inf Serv Use 2006;26(1):1-10 [FREE Full text] [Medline: 17060946]

Abbreviations

AIHW: Australian Institute of Health and Welfare
eHEALS: eHealth Literacy Scale
HISB: health information-seeking behaviors
ICC: intraclass correlation coefficient
PAM-13: 13-item Patient Activation Measure

Edited by G Eysenbach; submitted 12.02.15; peer-reviewed by S Medlock, J Chiu, M Silver; comments to author 05.08.15; revised version received 15.09.15; accepted 09.11.15; published 29.12.15.

Please cite as:
Lee K, Hoti K, Hughes JD, Emmerton LM
Consumer Use of “Dr Google”: A Survey on Health Information-Seeking Behaviors and Navigational Needs
J Med Internet Res 2015;17(12):e288
URL: http://www.jmir.org/2015/12/e288/
doi:10.2196/jmir.4345
PMID:26715363

Original Paper

General Practitioners' Concerns About Online Patient Feedback: Findings From a Descriptive Exploratory Qualitative Study in England

Salma Patel[1], BSc (Hons); Rebecca Cain[1], BA(Hons), PhD; Kevin Neailey[1], BSc (Hons), PhD; Lucy Hooberman[1], MA (Hons) Oxon, FRSA

[1]WMG, University of Warwick, Coventry, United Kingdom

Corresponding Author:
Salma Patel, BSc (Hons)
WMG
University of Warwick
International Digital Laboratory
Coventry, CV4 7AL
United Kingdom
Phone: 44 24 7657 5951
Fax: 44 24 7657 5951
Email: salma.patel@warwick.ac.uk

Abstract

Background: The growth in the volume of online patient feedback, including online patient ratings and comments, suggests that patients are embracing the opportunity to review online their experience of receiving health care. Very little is known about health care professionals' attitudes toward online patient feedback and whether health care professionals are comfortable with the public nature of the feedback.

Objective: The aim of the overall study was to explore and describe general practitioners' attitudes toward online patient feedback. This paper reports on the findings of one of the aims of the study, which was to explore and understand the concerns that general practitioners (GPs) in England have about online patient feedback. This could then be used to improve online patient feedback platforms and help to increase usage of online patient feedback by GPs and, by extension, their patients.

Methods: A descriptive qualitative approach using face-to-face semistructured interviews was used in this study. A topic guide was developed following a literature review and discussions with key stakeholders. GPs (N=20) were recruited from Cambridgeshire, London, and Northwest England through probability and snowball sampling. Interviews were transcribed verbatim and analyzed in NVivo using the framework method, a form of thematic analysis.

Results: Most participants in this study had concerns about online patient feedback. They questioned the validity of online patient feedback because of data and user biases and lack of representativeness, the usability of online patient feedback due to the feedback being anonymous, the transparency of online patient feedback because of the risk of false allegations and breaching confidentiality, and the resulting impact of all those factors on them, their professional practice, and their relationship with their patients.

Conclusions: The majority of GPs interviewed had reservations and concerns about online patient feedback and questioned its validity and usefulness among other things. Based on the findings from the study, recommendations for online patient feedback website providers in England are given. These include suggestions to make some specific changes to the platform and the need to promote online patient feedback more among both GPs and health care users, which may help to reduce some of the concerns raised by GPs about online patient feedback in this study.

(J Med Internet Res 2015;17(12):e276) doi:10.2196/jmir.4989

KEYWORDS

online reviews; physician quality; primary care; Internet; quality; patient empowerment; quality transparency; public reporting; attitude of health personnel; delivery of health care; feedback

Introduction

There has been growth in the use of online consumer feedback and review websites (eg, TripAdvisor), which some argue has allowed for transparent information and communication to influence change and has provided opportunities for consumers to read reviews and make more informed choices [1-3]. Similarly, there has been a growth in the volume of online patient ratings and comments, which suggests that patients in England (and other parts of the world) are embracing the opportunity to review their health care online [4-7]. There has also been a growth in the development of online patient feedback, with some patients now reviewing not just their experience of receiving health care, but also their medication and treatment plan [8].

Online patient feedback, in the context of this paper, can be defined as experiential feedback, ratings, reviews, and comments left by patients, carers, or service users on public Web-based platforms in England, such as NHS Choices, Patient Opinion, and iWantGreatCare, and on apps such as the iPhone-based Great Care app. Users can leave feedback and rate their experience with a general practitioner (GP) service, hospital, dentists, and other health care services, which is available in the public domain for other users to look at (see Figure 1 for an example of an online patient feedback website). The purpose of such websites is to give patients a voice by allowing them to leave feedback online, which some suggest will increase transparency, improve the quality of care, and could be used for service improvement [7,9]. Patients and carers could also then use these ratings and reviews to decide which health care provider to use; in England, this is part of the "patient choice" agenda [10-12].

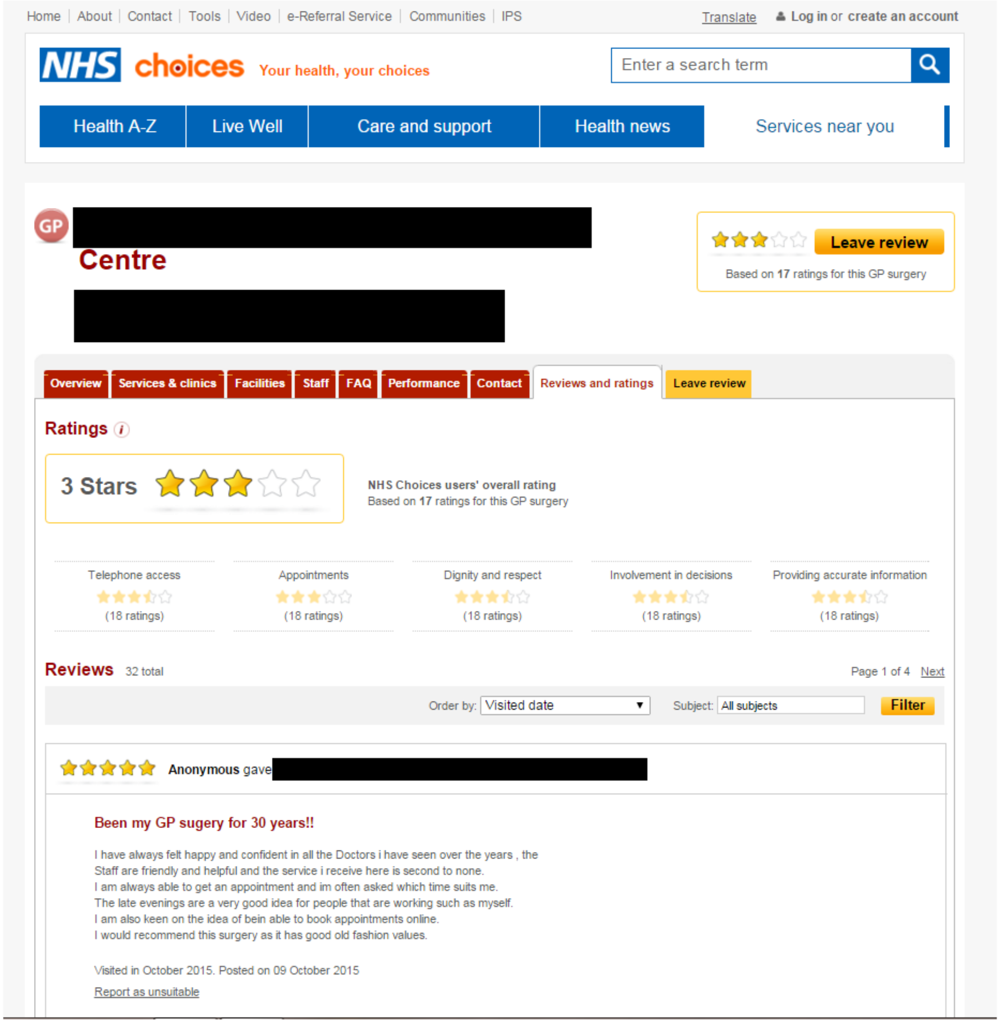
There are two major websites in the United Kingdom that (1) collect online reviews and ratings from patients about their experience of receiving care from their GP, and (2) allow the public to read patient feedback, which may be used by some patients or carers to choose a GP or a GP practice. The first is the NHS Choices website, which allows patients to leave comments under the GP practice's name, but does not allow an individual GP to be named or identified in any feedback [13]. The second is the iWantGreatCare website, which allows patients to leave feedback under a GP's name; therefore, GPs are named and identified in any review or feedback left on this website [14].

Research into online patient feedback or online physician-rating websites has been steadily increasing over the past few years with studies conducted in the United Kingdom [5-7,12,15-19], United States [20,21], Germany [22-26], and Australia [27] adding to the growing literature [28]. There appears to be some evidence to suggest that there is an association between online ratings and quality of care [4,5,21,29,30]. However, the results are not always consistent and, for some studies, the effect size is weak [28]. In particular, the extent to which the online ratings reflect the quality of primary care is less clear [6]. Furthermore, it is difficult to cross-apply the findings of different countries and different online patient feedback platforms because the characteristics of each online platform varies and the culture, context, and policies of each country and health care organization are different [9]. Despite this, some still argue that the "rich source of data" from online patient feedback has the potential to track quality of care [4,5,21,29,30].

Outside of the United Kingdom, a few studies have explored the type of patients who use online patient feedback platforms [24,31-34], whereas other studies have explored the attitudes of patients toward online patient feedback or doctor-rating websites [33,35,36]. A few physician representatives in the media have argued against the introduction of online feedback and rating websites by suggesting that they are dangerous and a waste of resources [37,38], and concerns about slander have also been raised by critics of such platforms [29,38-40]. However, there is very little research that explores health care professionals' attitudes toward online patient feedback.

Because health care professionals are one of the primary recipients of online patient feedback, one of the aims of this study was to explore and describe GPs' views about online patient feedback. This paper narrates the concerns raised by GPs in relation to online patient feedback only and the other findings, including the benefits of online patient feedback suggested by GPs, will be reported elsewhere. It is hoped that the findings from this study could be used to improve online patient feedback from the GPs' perspective and this may help to increase usage of online patient feedback by GPs and, by extension, their patients too.

Figure 1. Example of online patient feedback on NHS Choices website.



Been my GP sugery for 30 years!!

I have always felt happy and confident in all the Doctors i have seen over the years , the Staff are friendly and helpful and the service i receive here is second to none.
I am always able to get an appointment and im often asked which time suits me.
The late evenings are a very good idea for people that are working such as myself.
I am also keen on the idea of bein able to book appointments online.
I would recommend this surgery as it has good old fashion values.

Visited in October 2015. Posted on 09 October 2015

Methods

Data Collection

Because very little is known about GPs' perceptions of online patient feedback, there was a need for in-depth exploration to capture attitudinal and experiential data; therefore, a qualitative approach using semistructured interviews was best suited. A topic guide was developed following the guidance suggested by Bryman [41], Matthews and Ross [42], and Tracy [43] (see Multimedia Appendix 1 for a copy of the topic guide). A literature review was used as a basis for the topic guide as well as information from discussions with key stakeholders, such as the main lead at one of the online patient feedback website providers in the United Kingdom and 4 GPs. Further issues raised by participants during the interviews were also pursued and participants were encouraged to draw on experiences to illuminate their responses. Three vignettes were also developed (see Multimedia Appendix 2) following the guidance of Barter and Renold [44] and were used as prompts if the participant had not seen an online feedback review website before.

Sampling and Recruitment

Participants were recruited from Cambridgeshire, London, and Northwest England. A probability sampling approach was employed initially to ensure a wide range of characteristics of participants as recommended by Bryman [41]. However, despite using various strategies (all described in Table 1), only 6 participants were recruited using probability sampling. Other researchers who attempted to recruit GPs for research in the United Kingdom also reported immense difficulties [45-51]. Therefore, because of the limited response rate and the difficulties with recruiting sufficient GPs for this study using probability sampling, we resorted to using snowball sampling and 14 further GPs were recruited this way using various approaches (detailed in Table 1). In total, 20 GPs were interviewed for this study because at that point thematic saturation [52] had been reached.

Table 1. Recruitment strategies and number of general practitioners recruited for this study.

Recruitment strategy	Number of GPs recruited
Probability sampling	
Direct invitation to GP (postal invitations were sent to 25 practices in Cambridgeshire, which were then followed up by phone calls. A primary care research network [PCRN] also sent a letter on our behalf inviting and promoting the study to research active GPs and practices in Coventry)	1
Invitation through practice managers (phone calls were made to 25 GP practice managers in Cambridge and a follow-up email was also sent; 13 further GP practices were then contacted through phone and then fax)	2
Promoting study in email-based GP newsletters (the study was promoted in the following GP newsletters: Cambridgeshire NHS Newsletter, Lewisham Weekly Newsletter, Lambeth PCT newsletter, South NHS North West London Newsletter, Participate Autumn Magazine)	3
Snowball sampling	
Email to acquaintances with potential GP contacts	1
Twitter call out to acquaintances with potential GP contacts	0
GPs emailing their GP acquaintances	5
Medical doctors phone calling their GP acquaintances	8
Total	20

Study Interviews and Participants

The interviews were conducted at the location that was most convenient to the participant, with a preference given to the GP practice where the GP worked. However, some GPs preferred to be interviewed at their home outside of working hours and one GP was interviewed at a private meeting room. GPs were paid £80 for their participation.

All participants were sent the participant information sheet beforehand either through email or in the post and this contained the aims of the study among other things. Informed consent was obtained from all participants before the start of the interview. Interviews were recorded using two digital voice recorders. The study had full ethical approval from the Biomedical and Scientific Research Ethics Committee at the University of Warwick.

The descriptive characteristics of the 20 participants interviewed are shown in Table 2. Although 60% (12/20) of the GPs interviewed were between the ages of 30 and 34 years, they represented a variation in relation to duration of experience as GP, type of GP, and gender.

Table 2. Demographics and practice characteristics of participants (N=20).

Baseline characteristic	Frequency, n (%)
Age (years)	
25-29	1 (5)
30-34	12 (60)
35-39	3 (15)
40-44	1 (5)
45-49	1 (5)
50-54	1 (5)
55-59	1 (5)
Gender	
Male	12 (60)
Female	8 (40)
Type of GP	
Salaried	6 (30)
Partner	7 (35)
Senior partner	2 (10)
Lead	1 (5)
Locum	4 (20)
Years practicing as GP	
1-5	14 (70)
6-10	2 (10)
11-15	1 (5)
16-19	1 (5)
≥20	2 (10)

Data Preparation and Analysis

Interviews were transcribed verbatim and each transcript was double-checked for inaccuracies. The transcripts were then exported to NVivo software and analyzed using the framework method. This is a form of thematic analysis developed in the 1980s by researchers at the NatCen Social Research and has been used widely since then, both in policy research and other areas [53-56]. A thematic framework was created which was refined and then applied to all the data. Categories and themes were refined and defined until explanations were formed and thematic maps were produced. The analysis was conducted by the first author (SP) and the thematic framework and thematic maps were checked by all authors.

Results

In this interview-based study, participants were asked about their experience, usage, and attitudes (both positive and negative) toward online patient feedback. Because of the richness, depth, and breadth of the interview data, it was only possible to report the concerns about online patient feedback raised by participants in this paper. Other findings, such as those related to GPs' perceived benefits of online patient feedback or their attitudes toward social media, will be reported elsewhere. However, to place the GPs concerns into context, we believe it is important to understand participants' usage of and experience with online patient feedback as well as their overall impression of patients leaving feedback online about them. This is narrated in the subsequent section.

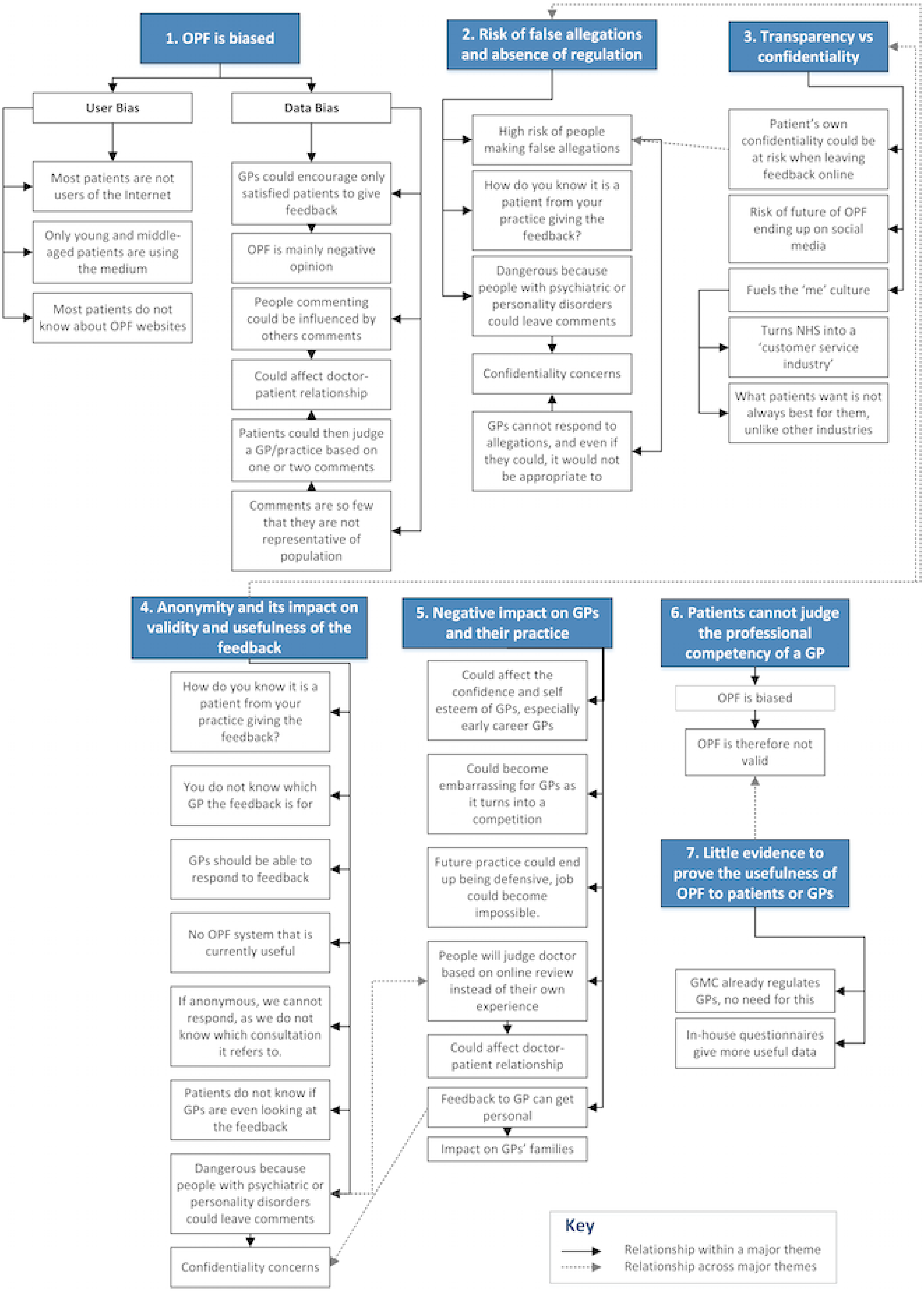
Usage, Awareness, and Overall Impression of Online Patient Feedback

Three-quarters of GPs interviewed in this study were aware that patients can leave feedback for them or their practice on the NHS Choices website. Four GPs had direct experience with online patient feedback and their practice or GPs in their practice had received feedback online on the NHS Choices website. One of the GPs also admitted he had received negative personal feedback online on the iWantGreatCare website. The majority of GPs interviewed (n=17) did not currently consider online patient feedback as a way of collecting feedback from patients. However, 12 participants believed that patients do have a right to place feedback about their GP online as long as the feedback was factually correct and on an appropriate website. Five participants, however, disagreed suggesting that patients do not have the right to place feedback about their GPs online.

Concerns About Online Patient Feedback

In this study, 56 individual concerns were raised by GPs (31 of which were unique) when asked the open-ended question: “Do you have any concerns about online patient feedback?” Other concerns raised about online patient feedback during the interview were also included in the analysis. Figure 2 is a thematic map that shows a summary of the concerns raised by participants and the 7 main themes (highlighted and numbered in the diagram) are discussed in the subsequent sections.

Figure 2. A thematic map illustrating the downsides of online patient feedback (OPF) according to the GP participants in this study.



Theme 1: Online Patient Feedback is Biased

User Bias

Most participants were concerned that it is only young and middle-aged patients who are leaving feedback on online patient feedback websites. Some participants commented that the majority of their patients were elderly and were certainly not using this medium. This, according to them, indicates that the feedback and ratings that are online are not representative of the overall experience of their patients and, therefore, they questioned its validity and usefulness:

> *You are getting your opinions from again, one particular sector of the community...their perceptions, their understanding and attitudes are different to the rest of the population...You have to use it with a bit of scrutiny there, in terms of interpreting the data, how do you put that in practice? If you were to put in any changes?* [P10]

One participant felt that there was not enough publicity about the NHS feedback website and many patients do not even know about the website.

Online Patient Feedback is Mainly Negative Opinion

One of the most repeated concerns raised by a quarter of participants was that online patient feedback is mainly negative opinion and is and will become a channel for disgruntled patients:

> *There's a bias towards putting negative feedback, if they've had a good experience, nonoutstanding one but an adequate one, they have no complaints but their needs are made [sic], they are unlikely to go on and post positive feedback.* [P13]

Other participants argued that it only takes one negative feedback to harm a GP's reputation:

> *One unfortunate comment or bit of a negative feedback, which may be taken well out of context, can harm your reputation.* [P16]

Online Patient Feedback is Too Small in Number

A few participants (n=3) raised concerns that feedback left for a GP or a GP practice on online websites is too small in number and, therefore, it is not representative of their record of performance:

> *It's a small number of comments, we've had 2 [online reviews] out of a population of 12,000 [on the NHS Choices website], I don't think that would be representative enough for a potential patient to go on and go, "alright ok they've got 50% bad comments, right I'm not registering there" (laugh).* [P2]

GPs were concerned that patients could judge a practice or GP based on a very small number of reviews (and make an invalid "choice") and this could also affect the doctor-patient relationship.

Reviews Could Be "Gamed"

Some participants (n=5) were concerned that reviews could be manipulated and that some GPs could encourage only satisfied patients to give feedback, which again would add bias to the data. Moreover, 4 participants from this study admitted they would only encourage those patients who they know will give positive feedback to leave feedback online:

> *No, obviously, if somebody has had a good experience, you might encourage it [leave feedback on NHS Choices]. But also, I think if somebody wants to make a complaint I would say you can write to the practice manager and there's always, I may not actively promote it [giving feedback on NHS Choices].* [P7]

Theme 2: Risk of False Allegations and Absence of Regulation

Half of the participants felt that there was a very high risk of patients leaving false allegations about them or the practice on online patient feedback websites. Furthermore, a quarter of participants felt that the owners of such websites (eg, the NHS Choices feedback site) were not regulating feedback left on these websites and removing malicious or factually incorrect comments from patients. Participants were particularly concerned that their patients with psychiatric or personality disorders could leave factually incorrect or malicious comments about them and harm their reputation:

> *You will have everyone, including people with severe psychiatric illness [leaving feedback on online patient feedback websites]...so I think it's [online patient feedback] potentially quite [a] dangerous tool.* [P19]

Also, a few participants (n=3) felt that even if GPs could respond online to such allegations, it would not be appropriate for them to respond online.

Theme 3: Transparency Versus Confidentiality

Eight participants agreed that patient reviews left online will seemingly help to increase transparency of care and improve the quality of care, and they were not concerned about the feedback being online and being so transparent as long as there was a "proper system" in place for online patient feedback:

> *It worries me if it's [online patient feedback] not a proper system.* [P18]

By "proper system," participants meant that the website was well regulated and validated. The website could verify, for example, that the patient leaving the feedback was an actual patient of that particular GP or that the patient did not have a malicious agenda. However, 8 participants were concerned about the platform being "too open" (P11) and in public due to the possibility of people making false allegations, and its damaging impact on the reputation and career of a GP and a GP's own personal confidentiality:

> *If it's [feedback] in public, particularly if I felt it was untrue...if you got y'know someone made an allegation...if that happened to a doctor it could*

> *destroy their career, and their self-esteem, and I just think that's not fair on doctors.* [P1]

A few participants felt that these types of websites would fuel what they called the "me" culture and turn the NHS into a "customer service industry" and were concerned that it could lead to patients in the future thinking it was perfectly fine to leave feedback about physicians on social media, where according to them, it is impossible to validate or regulate the feedback. Another participant went on to explain that, particularly in health care, what a patient wants is not always what is best for them:

> *There can be a difference certainly between what people need and what people want, and if people don't get what they want, often they can feedback negatively about that, even though actually the doctor or the medical provider or whatever who is looking after them, has done exactly the right thing.* [P2]

In addition to worrying about the confidentiality of GPs themselves, a few participants were also concerned that the transparent nature of the feedback meant that a patient's own confidentiality may be at risk because they may feel the need to disclose personal health information about themselves on a public platform. Some participants (n=3) were also concerned that GPs are unable to respond to patient reviews online due to the possibility of violating doctor-patient confidentiality because they may need to disclose health information about the patient in their response.

Theme 4: Anonymity and its Impact on Validity and Usefulness of the Feedback

All participants (N=20) were aware that the feedback left on NHS Choices was left anonymously by patients. Some participants raised concerns that because the feedback was left anonymously, they would not know which consultation it referred to and, therefore, could not respond to the feedback nor make real use of it for improvement. Others (n=5) felt that the anonymous nature of the feedback meant they would not know if it was an actual patient from the practice that had left the feedback and questioned whether such feedback is even valid:

> *Again, if it is anonymous, then, with any feedback really, you really don't know, is it somebody from this practice, or somebody, well it could be anybody really leaving a feedback there [on online patient feedback websites].* [P10]

Participants were then asked specifically if the feedback would be more useful to them if it was not left anonymously. Seven participants said that it would be more useful to them if feedback was left with the patient's real name so that they can then look up the consultation and see what went wrong:

> *If you had their name there, you could obviously understand where this is coming from, and then you can think about it or go back on it, and make ways of improving yourself during your consultation skills. But if it is very anonymous...out of seeing 40 patients in a day, 200 in a week, which one are we talking about, in terms of who?* [P10]

However, the remaining 13 participants disagreed, commenting that it would not be fair on patients to give their real name because, according to them, it will affect the doctor-patient relationship and patients will not leave feedback online if they cannot leave it anonymously. One participant appeared to suggest a solution that patients should leave their NHS number when they leave feedback to verify that they are a patient registered at that particular practice. Another participant raised the question that despite patients not naming themselves when leaving feedback online, would patients really remain anonymous because sometimes it was easy to identify a particular patient from an anonymous online comment.

Theme 5: Negative Impact on General Practitioners and Their Practice

In addition to the threat of defamation discussed previously and its impact on GP reputation and career, 2 participants were also concerned that negative feedback online could affect the self-confidence and self-esteem of GPs, which would in turn affect their practice, especially those GPs who are early in their career:

> *It [online patient feedback] will affect people in their early career a lot more, and could break their confidence and make them insular. Is that what you really want to be doing to your future doctors?* [P11]

Some participants also felt that people will start judging GPs based on online reviews instead of their own experience and this could also affect not only the doctor-patient relationship, but also their practice. Furthermore, participants raised concerns that due to the possibility of negative reviews going online (whether true or false), future practice could end up being defensive and it would be impossible to practice properly:

> *I don't know how on earth we are going to have a decent relationship...doctors have become so defensive already...just to make sure they don't get things online, or do you want them to actually do right for you...give you good care in the right manner in the right timeframe, in a manner which is satisfactory to, or do want them to just do things because they are so scared of litigation of online feedback.* [P11]

Other GPs raised concerns that it could become embarrassing for them if their practice became public and turned into a "competition" and this could impact patient care too. One participant was particularly concerned about the negative impact online patient feedback could have on her family:

> *I suppose it's just the fact that something that's online...you think about your family and other people, close to you nearest and dearest, sort of looking at things and getting upset on your behalf as well.* [P12]

Theme 6: Patients Cannot Judge the Professional Competence of a General Practitioner

Some of the participants who were not in favor of online patient feedback argued that the General Medical Council was already regulating them, so there was no need for patients attempting to "regulate" them online and, in fact, how can patients judge whether a GP was competent or not?

> *Can you really say a patient has that ability to say whether you are underperforming or not?...so the people that are doing appraisal and revalidation are also GPs, they know what you should be doing. I think they should police it, as opposed to patients.* [P14]

Theme 7: Little Evidence to Prove Usefulness of Online Patient Feedback to Patients or General Practitioners

Two participants argued that there was no evidence currently to prove the usefulness of online patient feedback to patients or GPs:

> *I think some things with Government policy or in the NHS policy are brought in without having any evidence of benefit, sometimes people jump at the chance "oh we will do this" and they don't think why.* [P1]

Furthermore, a few participants argued that existing methods of collecting patient feedback, such as in-house questionnaires, were perfectly adequate and gave more useful data. However, when asked separately about offline feedback, more than half of the participants commented that they did not collect "useful" data.

Discussion

To our knowledge, this study is the first study that explores GPs' concerns about online patient feedback. The study's findings suggest that GPs have reservations about online patient feedback and question online patient feedback's validity, value, usability, and its transparent nature, and are worried about how this will impact them, their practice, and their patients.

Validity and Value of Online Patient Feedback

In this study, GPs were concerned—among other things—about the bias of online patient feedback, both from the user and data perspective. The concern of potential bias due to the age of patients using online patient feedback (in favor of younger patients) has also been raised in literature by some [39,57-59] and some studies appear to support it [6,31,33,36]. However, it could be argued that even patient satisfaction results that are offline are influenced by age, education, and health status [60]. Furthermore, Greaves et al [5] argue that although there may be risks in using ratings from a small group of self-selecting patients, according to them it is outweighed by its positives, mainly that online patient feedback is low cost and has the ability to detect episodes of poor care that a traditional survey may miss. However, as some participants highlighted in this study, this does not address the question of whether a rating of a particular GP or a GP practice can truly be representative, valid, and fair if only the younger and middle-aged patients are leaving ratings or feedback. This is especially crucial for those practices and GPs that serve a largely elderly population.

Furthermore, other participants argued that online patient feedback is mainly negative opinion and will become a channel for disgruntled patients. This sentiment has been raised in opinion articles [37,39] and literature [29,61], but has been counteracted by the argument that many studies (including [6,19,20,22,29,34,62-66]) have found that the majority of feedback left on online physician review websites is actually positive [28]. However, Greaves et al [6] found that the recommendation level of GP and practices in England for the same period was 64% online and 82% in patient surveys. This, they suggest, does indicate that there may be a selection bias in online patient feedback toward less satisfied patients versus when patients are selected randomly and this appears to suggest that the concerns raised by participants in this study may be valid. Furthermore, Merrell et al [61] argue that the abundance of positive reviews cannot negate the impact of negative ones because negative ones, however few, can have long-lasting ramifications as a few GPs in this study highlighted. This is also supported by findings from a study by Hanauer et al [67] who found that parents who are exposed to a positive recommendation of a physician from a neighbor are less likely to choose that physician for their child if they were then exposed to negative reviews about that same physician online. However, Adams [68] found that patient reviews online are not always inherently positive or negative; rather, they contain a mixture of positive and negative comments as well as references to and comparisons with previous health care experiences.

Another argument put forward by GPs in this study was that patients cannot judge the professional competence of a GP; therefore, how can online patient feedback be a true representation of their practice? The concern of whether a patient can adequately judge quality of care received was also raised by Lagu and Lindenauer [58]. It would be useful to explore in future research whether patients are aware that patient-led ratings may be based primarily on the bedside manner of a GP, according to some GPs in this study, and not necessarily on the clinical competence of a GP.

The GPs interviewed were also concerned that feedback left online for them are too few in number and therefore not representative of their performance. This appears to be supported by a study on the NHS Choices feedback website, which found that only 61% of GP practices in England had been reviewed and the number of ratings left per practice was variable with an average of only 2 ratings per practice [6]. However, this study by Greaves et al [6] explored data from more than 5 years ago (between October 2009 and December 2010) and more up-to-date analysis of such websites is required to truly understand the current state because usage may have changed. Taking into account that reviews on the NHS website only correspond to 0.005% of all GP consultations [6] and that studies from the United States [29,66,69,70], Germany [34,71,72], and Australia [27] all indicate that less than 30% of physicians have been rated (and even those that have been rated have on average less than 4 ratings each), the assertions raised by the participants in this study may be true and valid, and need to be addressed by online patient feedback platform providers.

Strech [40] suggests that the solution to this may be that ratings should not be made available until they reach a certain baseline number (eg, 5-10). Although individual pieces of feedback could be displayed before a baseline number is reached, the overall star rating, for example, should not be shown until there are a reasonable number of ratings left for a practice or GP. If the NHS and other online patient feedback website providers want GPs to take these reviews seriously and for the ratings per practice to be “valid” and representative (so that patients can make an accurate “choice”), it needs to do more to get patients to leave reviews (see Textbox 1 for a list of recommendations for online patient feedback providers based on this study). Although the overall rating may not be representative of the quality of care provided by GPs at a GP practice, this does not mean that the individual patient feedback left, however few, may not be useful to GPs and practices to use to make changes and identify opportunities for improvement. This suggests that even if some feedback providers may choose not to publish reviews until a certain baseline has been reached, the unpublished reviews could be sent to GPs to review and use for improvement.

Textbox 1. Recommendations for online patient feedback website providers based on findings from this study.

Based on the findings from this study with GPs, recommendations for online patient feedback website providers in England are as follows:

1. Promote online patient feedback more among GPs and patients

- Promote online patient feedback among GPs and train GPs to use online patient feedback. This will help to reduce misunderstandings about online patient feedback among GPs, which may help to increase usage of online patient feedback by GPs and also by patients. If GPs believe online patient feedback is valid and useful, they are more likely to promote it to their patients and this may be one of the most effective ways to promote the platform with patients. This could be done through training or may even be as simple as creating a document entitled “A Guide to Online Patient Feedback [specific platform name] for GPs” and signpost it well, both online and offline.
- Implement a campaign to promote online patient feedback to patients. This will help to increase the number of patients and type of patients leaving feedback and reviews and therefore the feedback left online is less likely to be biased and unrepresentative. This may mean GPs will take it more seriously and patients will be able to make a more valid “choice.” This could be done through traditional marketing routes through GP practices and digital methods, such as social media and TV ads.

2. Convince and reassure GPs about the value of online patient feedback

- Outline precisely how feedback left on the website is moderated and regulated, especially in relation to malicious or personal comments about individual GPs.
- Outline on the website and in any marketing leaflets what GPs can do with feedback that is left online for them, in particular, how to respond to it and use it for improvement.
- Make patients aware that feedback and ratings left by other patients on these online patient feedback websites may be based primarily on bedside manner and that the majority of patients do not have the ability to judge the professional competence of a GP.

3. Consider some changes to the online patient feedback website

- To eliminate concerns about patients judging a GP or a practice based on just a handful of reviews, have a larger number of reviews on the website per practice before the overall rating is calculated and shown.
- Validate that the patient leaving feedback on the online patient feedback website is registered as a patient at the given practice through, for example, asking the patient for his/her NHS number. The NHS number could be concealed from the practice to protect the identity of the patient.
- Allow patients to leave feedback both for individual GPs and for the practice.
- Create an aggregated score of results of measures of competence of GP and patient feedback and reviews, left online and offline, instead of the rating being based on just a few reviews left online at the moment. This has been recommended by the Nuffield Trust [73].

The limited number of online reviews for GPs and practices may be partly explained by one participant’s comment that “patients do not even know about online patient feedback.” This appears to be supported by a study that found that only 15% of the 200 participants in one borough of London were aware of the existence of online feedback websites in health care [36]. However, this study was conducted almost 3 years ago and the awareness of online patient feedback among Londoners may have increased. Nevertheless, there is little evidence as to what extent the NHS Choices feedback website is known and used by patients in the United Kingdom. However, in the United States and Germany, recent studies found that approximately a quarter of respondents had used a physician-rating website [24,33]. This may be partly due to the higher usage and popularity of private health care in both the United States and Germany.

Even when reviews or feedback are left for GPs, some GPs in this study were concerned that the feedback or reviews left by patients could be manipulated without the GP or the practice doing anything “illegal” and this could add serious bias to the data and it would question the validity of the overall rating. This concern is similar to concerns raised in literature that ratings could be “gamed” by organizations or individuals and people could leave fake or multiple entries [6,58,74]. Lagu et al [66] analyzed feedback on review websites in the United States and found several reviews they felt had been written by the physician because they contained information only the physician would know. In another study, Kadry et al [63] found some reviews that they believed were acts of sabotage from competing providers.

A few other GPs in this study argued that there was not enough evidence to prove the usefulness of online patient feedback to patients or GPs. Although research into online physician-rating websites has been steadily increasing over the past few years and studies conducted in the United Kingdom, United States, Germany, the Netherlands, and Australia are all adding to the growing literature [4,5,21,29,30], there is currently a huge gap in the literature. For example, further research is needed to determine whether patients believe online patient feedback is "useful" to them to give feedback or to use to choose a health care provider.

Transparency of Online Patient Feedback

The transparent nature of online feedback websites is what has made it so attractive to patients and health policy makers because the understanding is that reviews left online will apparently increase transparency of care and improve the quality of care [63,68]. A few participants in this study appeared to support this view, whereas the majority had concerns about the platform being "too open" (P11) and in public due to the high risk of people making false allegations and its damaging impact on the reputation and career of a GP, a GP's self-confidence, self-esteem, and personal confidentiality, all of which could affect their professional practice. Concerns about slander have also been raised by critics of online patient feedback platforms, mainly physician representatives such as the British Medical Association [29,38-40]. However, NHS Choices in England claim to have a strict set of regulations that they use to protect physicians and hospitals from content that may damage their reputation [13]. Despite the NHS Choices promising that all "inflammatory remarks" are removed, it is unknown how this is put into practice and to what extent, and also what constitutes "inflammatory." Owners of such websites need to make this clear to their users [40]. Furthermore, a few participants in this study remarked that although NHS Choices may anonymize the doctor, it was easy for GPs and the public to work out which GP or staff member the comment was directed to; therefore, it does not really give them the anonymity and protection it claims to.

Usability of Online Patient Feedback

Others remarked that due to the GP being anonymous in comments, it was difficult to work out who the comment was for and, therefore, could not be used for improvement. This concern was also raised by McCartney [39] who as a practicing GP felt that it was difficult, if not impossible, for doctors to learn from anonymous comments. One participant explained that in his opinion the difference is related to the size of the practice; where there are fewer GPs in a practice, it is easy to work out who the feedback is for and their reputation could be harmed much more easily. Another participant felt that harming of reputation was happening offline too, so it made no difference whether it was online or offline. However, others remarked that being online was "too public" and hundreds and thousands of people could have access to it. Some physicians have gone as far as getting a court order to remove an online review according to Kadry et al [63], but they argue that it is very difficult to defend against online misinformation and defamation. Further research is required to determine how patients feel about remaining anonymous and naming their GP online when leaving feedback about their GP online and whether remaining anonymous and naming their GP are key criteria for them to leave feedback on online patient feedback websites.

Conclusions and Recommendations

The majority of GPs interviewed in this study had concerns and reservations about online patient feedback because they felt that online patient feedback was not an accurate representation of their performance due to user bias and data bias. They were also worried about the impact this could have on them, on their practice, and their patients, who may use these "questionable online ratings" to make an "invalid choice" of which health care provider to use. GPs in this study also felt that due to the transparent nature of the feedback online and what they perceive to be lack of regulation, there is a high risk of false allegations being left about them, which could have an impact on them personally, on their family, on their professional practice (more defensive medicine), and on their relationship with their patients. Other GPs questioned the usefulness of the online feedback if the feedback is left anonymously, but acknowledged the benefits to patients of leaving feedback anonymously. A few participants also argued that there was no current evidence to prove online patient feedback's usefulness to GPs or patients.

Our findings suggest that most concerns raised by GPs may be valid and need to be addressed by online patient feedback providers and other online patient feedback stakeholders. If the NHS and other online patient feedback website providers want GPs to take these reviews seriously, for example, and for the ratings per practice to be valid and representative (so that patients can make an accurate "choice"), they need to do more to get patients to leave reviews. Promoting online patient feedback among GPs and reassuring them of the safety and usefulness of such platforms may also mean GPs are more likely to use online patient feedback for their own professional development and encourage their patients to leave feedback on online patient feedback websites. Other recommendations for online patient feedback website providers based on findings from this study can be found in Textbox 1.

Limitations of the Study

One of the aims of this descriptive study was to explore GPs' concerns about online patient feedback and the qualitative findings from this study were not intended to be representative of all GPs in England. We do acknowledge that the sample size for this study was small (N=20) and because 60% of participants were between the ages of 30 and 34 years, there may have been a sample bias toward more technology-savvy GPs. However, we found little difference in Internet usage of all the different-aged participants in our sample. We did attempt to recruit participants randomly to get GPs from different age groups and backgrounds, and more than a quarter of our participants were recruited using probability sampling.

Despite our findings not being representative of all GPs in England and this paper being limited to narrating GPs' negative attitudes to online patient feedback only, the findings highlight key concerns related to online patient feedback from GPs' perspective and place them into the context of existing literature and viewpoints. This helped form recommendations for feedback providers and can help inform further research in this area.

Acknowledgments

This research was funded by the Engineering and Physical Sciences Research Council (EPSRC) under the Participation in Healthcare Environment Engineering Programme (#EP/H022031/1). Thank you to the 20 GPs who gave their time to contribute their experiences and viewpoints in interviews. A special note of thanks to those who assisted with recruiting GPs to the study, in particular, Dr Imran Satia, Dr Hamed Khan, and the South Midlands PCRN team.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A copy of the complete topic guide used for the interviews in this study.

[PDF File (Adobe PDF File), 24KB - jmir_v17i12e276_app1.pdf]

Multimedia Appendix 2

A copy of the vignettes used during interviews with the GPs.

[PDF File (Adobe PDF File), 132KB - jmir_v17i12e276_app2.pdf]

References

1. Schuckert M, Liu X, Law R. Insights into suspicious online ratings: direct evidence from TripAdvisor. Asia Pacific Journal of Tourism Research 2015 Apr 21:1-14 [FREE Full text] [doi: 10.1080/10941665.2015.1029954]
2. Li H, Ye Q, Law R. Determinants of customer satisfaction in the hotel industry: an application of online review analysis. Asia Pacific Journal of Tourism Research 2013 Oct 25;18(7):784-802. [doi: 10.1080/10941665.2012.708351]
3. Filieri R. What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. Journal of Business Research 2015 Jun;68(6):1261-1270. [doi: 10.1016/j.jbusres.2014.11.006]
4. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]
5. Greaves F, Pape U, King D, Darzi A, Majeed A, Wachter R, et al. Associations between Internet-based patient ratings and conventional surveys of patient experience in the English NHS: an observational study. BMJ Qual Saf 2012 Jul;21(7):600-605. [doi: 10.1136/bmjqs-2012-000906] [Medline: 22523318]
6. Greaves F, Pape U, Lee H, Smith D, Darzi A, Majeed A, et al. Patients' ratings of family physician practices on the internet: usage and associations with conventional measures of quality in the English National Health Service. J Med Internet Res 2012;14(5):e146 [FREE Full text] [doi: 10.2196/jmir.2280] [Medline: 23076301]
7. Trigg L. Patients' opinions of health care providers for supporting choice and quality improvement. J Health Serv Res Policy 2011 Apr;16(2):102-107. [doi: 10.1258/jhsrp.2010.010010] [Medline: 20870690]
8. Hughes S, Cohen D. Can online consumers contribute to drug knowledge? A mixed-methods comparison of consumer-generated and professionally controlled psychotropic medication information on the internet. J Med Internet Res 2011;13(3):e53 [FREE Full text] [doi: 10.2196/jmir.1716] [Medline: 21807607]
9. Greaves F, Millett C, Nuki P. England's Experience incorporating "anecdotal" reports from consumers into their national reporting system: lessons for the United States of what to do or not to do? Med Care Res Rev 2014 Oct;71(5 Suppl):65S-80S. [doi: 10.1177/1077558714535470] [Medline: 24836765]
10. Fotaki M. What Market-Based Patient Choice Can’t Do for the NHS: The Theory and Evidence of How Choice Works in Health Care. London: Centre for Health and the Public Interest; 2014 Mar. URL: http://chpi.org.uk/wp-content/uploads/2014/03/What-market-based-patient-choice-cant-do-for-the-NHS-CHPI.pdf [accessed 2015-11-07] [WebCite Cache ID 6crZO3XgL]
11. House of Commons Library. 2015 Feb 03. General practice in England today URL: http://www.parliament.uk/mps-lords-and-offices/offices/commons/commonslibrary/ [accessed 2015-11-30] [WebCite Cache ID 6dQTkJ7l2]
12. King D, Zaman S, Zaman S, Kahlon G, Naik A, Jessel A, et al. Identifying quality indicators used by patients to choose secondary health care providers: a mixed methods approach. JMIR Mhealth Uhealth 2015;3(2):e65 [FREE Full text] [doi: 10.2196/mhealth.3808] [Medline: 26048441]
13. NHS Choices. 2014 Sep 18. Our terms, conditions and policies: comments policy URL: http://www.nhs.uk/aboutNHSChoices/aboutnhschoices/termsandconditions/Pages/commentspolicy.aspx [accessed 2015-07-30] [WebCite Cache ID 6aPjI9Rnb]
14. iWantGreatCare. URL: https://www.iwantgreatcare.org/ [accessed 2015-10-16] [WebCite Cache ID 6cJkcgzv9]
15. Greaves F, Millett C. Consistently increasing numbers of online ratings of healthcare in England. J Med Internet Res 2012;14(3):e94 [FREE Full text] [doi: 10.2196/jmir.2157] [Medline: 22742977]

16. Greaves F, Laverty A, Ramirez-Cano D, Pulman S, Moilanen K, Darzi A, et al. Analysis of patients' comments about hospitals in the English NHS via Twitter, and comparison with patient surveys. BMJ Quality & Safety 2014 Mar 17;23(4):348-348 [FREE Full text] [doi: 10.1136/bmjqs-2014-002893.5]
17. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]
18. Greaves F, Pape U, King D, Darzi A, Majeed A, Wachter RM, et al. Associations between Web-based patient ratings and objective measures of hospital quality. Arch Intern Med 2012 Mar 12;172(5):435-436. [doi: 10.1001/archinternmed.2011.1675] [Medline: 22331980]
19. Lagu T, Goff SL, Hannon NS, Shatz A, Lindenauer PK. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. Jt Comm J Qual Patient Saf 2013 Jan;39(1):7-15. [Medline: 23367647]
20. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012 Jun;27(6):685-692 [FREE Full text] [doi: 10.1007/s11606-011-1958-4] [Medline: 22215270]
21. Bardach N, Asteria-Peñaloza R, Boscardin W, Dudley R. The relationship between commercial website ratings and traditional hospital performance measures in the USA. BMJ Qual Saf 2013 Mar;22(3):194-202 [FREE Full text] [doi: 10.1136/bmjqs-2012-001360] [Medline: 23178860]
22. Emmert M, Meier F, Heider A, Dürr C, Sander U. What do patients say about their physicians? an analysis of 3000 narrative comments posted on a German physician rating website. Health Policy 2014 Oct;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]
23. Emmert M, Sander U, Esslinger AS, Maryschok M, Schöffski O. Public reporting in Germany: the content of physician rating websites. Methods Inf Med 2012;51(2):112-120. [doi: 10.3414/ME11-01-0045] [Medline: 22101427]
24. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. J Med Internet Res 2013;15(8):e187 [FREE Full text] [doi: 10.2196/jmir.2702] [Medline: 23985220]
25. Emmert M, Halling F, Meier F. Evaluations of dentists on a German physician rating Website: an analysis of the ratings. J Med Internet Res 2015;17(1):e15 [FREE Full text] [doi: 10.2196/jmir.3830] [Medline: 25582914]
26. Emmert M, Sander U, Pisch F. Eight questions about physician-rating websites: a systematic review. J Med Internet Res 2013;15(2):e24 [FREE Full text] [doi: 10.2196/jmir.2360] [Medline: 23372115]
27. Atkinson S. Current status of online rating of Australian doctors. Aust J Prim Health 2014;20(3):222-223. [doi: 10.1071/PY14056] [Medline: 24852125]
28. Verhoef L, Van de Belt TH, Engelen LJ, Schoonhoven L, Kool R. Social media and rating sites as tools to understanding quality of care: a scoping review. J Med Internet Res 2014;16(2):e56 [FREE Full text] [doi: 10.2196/jmir.3024] [Medline: 24566844]
29. Gao G, McCullough J, Agarwal R, Jha A. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. J Med Internet Res 2012;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]
30. Segal J, Sacopulos M, Sheets V, Thurston I, Brooks K, Puccia R. Online doctor reviews: do they track surgeon volume, a proxy for quality of care? J Med Internet Res 2012;14(2):e50 [FREE Full text] [doi: 10.2196/jmir.2005] [Medline: 22491423]
31. Bidmon S, Terlutter R, Röttl J. What explains usage of mobile physician-rating apps? Results from a web-based questionnaire. J Med Internet Res 2014;16(6):e148 [FREE Full text] [doi: 10.2196/jmir.3122] [Medline: 24918859]
32. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Public awareness, perception, and use of online physician rating sites. JAMA 2014 Feb 19;311(7):734-735. [doi: 10.1001/jama.2013.283194] [Medline: 24549555]
33. Terlutter R, Bidmon S, Röttl J. Who uses physician-rating websites? Differences in sociodemographic variables, psychographic variables, and health status of users and nonusers of physician-rating websites. J Med Internet Res 2014;16(3):e97 [FREE Full text] [doi: 10.2196/jmir.3145] [Medline: 24686918]
34. Emmert M, Meier F. An analysis of online evaluations on a physician rating website: evidence from a German public reporting instrument. J Med Internet Res 2013;15(8):e157 [FREE Full text] [doi: 10.2196/jmir.2655] [Medline: 23919987]
35. The Tavistock Institute. Potential benefits of web-based feedback for improving health service quality: Lessons emerging from the evaluation of Patient Opinion website Internet URL: http://www.tavinstitute.org/lectures_and_presentations/powerpoint/potential-benefits-of-web-based-feedback-for-improving-health-service-quality-lessons-emerging-from-the-evaluation-of-patient-opinion-website/ [accessed 2012-01-23] [WebCite Cache ID 6crZZiRJp]
36. Galizzi MM, Miraldo M, Stavropoulou C, Desai M, Jayatunga W, Joshi M, et al. Who is more likely to use doctor-rating websites, and why? A cross-sectional study in London. BMJ Open 2012;2(6):N/A-N/A [FREE Full text] [doi: 10.1136/bmjopen-2012-001493] [Medline: 23148340]
37. Boffey D. The Guardian. 2011 Dec 03. GPs' fury over "poison pen" online tirades URL: http://www.theguardian.com/society/2011/dec/03/gps-poison-pen-online [accessed 2015-11-07] [WebCite Cache ID 6crZfvo3C]

38. Todd R. Digital Health. 2012 Mar 03. GPs give low rating to Choices comments URL: http://www.digitalhealth.net/news/27415/gps-give-low-rating-to-choices-comments [accessed 2015-11-07] [WebCite Cache ID 6crZlDwkj]
39. McCartney M. Will doctor rating sites improve the quality of care? No. BMJ 2009;338:b1033. [Medline: 19293224]
40. Strech D. Ethical principles for physician rating sites. J Med Internet Res 2011;13(4):e113 [FREE Full text] [doi: 10.2196/jmir.1899] [Medline: 22146737]
41. Bryman A. Social Research Methods. Oxford: Oxford University Press; 2008.
42. Matthews B, Ross L. Research Methods: A Practical Guide for the Social Sciences. Harlow: Pearson Education Limited; 2010.
43. Tracy S. Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact. Chichester: Wiley-Blackwell; 2013.
44. Barter C, Renold E. 'I wanna tell you a story': exploring the application of vignettes in qualitative research with children and young people. International Journal of Social Research Methodology 2000 Jan;3(4):307-323 [FREE Full text] [doi: 10.1080/13645570050178594]
45. Hummers-Pradier E, Scheidt-Nave C, Martin H, Heinemann S, Kochen MM, Himmel W. Simply no time? Barriers to GPs' participation in primary health care research. Fam Pract 2008 Apr;25(2):105-112 [FREE Full text] [doi: 10.1093/fampra/cmn015] [Medline: 18417465]
46. Williamson M, Pirkis J, Pfaff J, Tyson O, Sim M, Kerse N, et al. Recruiting and retaining GPs and patients in intervention studies: the DEPS-GP project as a case study. BMC Med Res Methodol 2007;7:42 [FREE Full text] [doi: 10.1186/1471-2288-7-42] [Medline: 17875219]
47. Shelton BJ, Wofford JL, Gosselink CA, McClatchey MW, Brekke K, Conry C, et al. Recruitment and retention of physicians for primary care research. J Community Health 2002 Apr;27(2):79-89. [Medline: 11936759]
48. Wetzel D, Himmel W, Heidenreich R, Hummers-Pradier E, Kochen M, Rogausch A, et al. Participation in a quality of care study and consequences for generalizability of general practice research. Fam Pract 2005 Aug;22(4):458-464 [FREE Full text] [doi: 10.1093/fampra/cmi022] [Medline: 15814583]
49. Templeton L, Deehan A, Taylor C, Drummond C, Strang J. Surveying general practitioners: does a low response rate matter? Br J Gen Pract 1997 Feb;47(415):91-94 [FREE Full text] [Medline: 9101692]
50. Barclay S, Todd C, Finlay I, Grande G, Wyatt P. Not another questionnaire! Maximizing the response rate, predicting non-response and assessing non-response bias in postal questionnaire studies of GPs. Fam Pract 2002 Feb;19(1):105-111 [FREE Full text] [Medline: 11818359]
51. Asch S, Connor S, Hamilton E, Fox S. Problems in recruiting community-based physicians for health services research. J Gen Intern Med 2000 Aug;15(8):591-599 [FREE Full text] [Medline: 10940152]
52. O'Reilly M, Parker N. 'Unsatisfactory Saturation': a critical exploration of the notion of saturated sample sizes in qualitative research. Qualitative Research 2012 May 17;13(2):190-197 [FREE Full text] [doi: 10.1177/1468794112446106]
53. Ritchie J, Lewis J, Nicholls CM, Ormston R. Qualitative Research Practice: A Guide for Social Science Students and Researchers. Thousand Oaks, CA: Sage Publications; 2013.
54. Dixon-Woods M. Using framework-based synthesis for conducting reviews of qualitative studies. BMC Med 2011;9:39 [FREE Full text] [doi: 10.1186/1741-7015-9-39] [Medline: 21492447]
55. Smith J, Firth J. Qualitative data analysis: the framework approach. Nurse Res 2011;18(2):52-62. [doi: 10.7748/nr2011.01.18.2.52.c8284] [Medline: 21319484]
56. Gale N, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
57. Rozenblum R, Bates DW. Patient-centred healthcare, social media and the internet: the perfect storm? BMJ Qual Saf 2013 Mar;22(3):183-186. [doi: 10.1136/bmjqs-2012-001744] [Medline: 23378660]
58. Lagu T, Lindenauer PK. Putting the public back in public reporting of health care quality. JAMA 2010 Oct 20;304(15):1711-1712. [doi: 10.1001/jama.2010.1499] [Medline: 20959582]
59. Tritter JQ, McCallum A. The snakes and ladders of user involvement: moving beyond Arnstein. Health Policy 2006 Apr;76(2):156-168. [doi: 10.1016/j.healthpol.2005.05.008] [Medline: 16006004]
60. Hekkert KD, Cihangir S, Kleefstra SM, van den Berg B, Kool RB. Patient satisfaction revisited: a multilevel approach. Soc Sci Med 2009 Jul;69(1):68-75. [doi: 10.1016/j.socscimed.2009.04.016] [Medline: 19446942]
61. Merrell JG, Levy BH, Johnson DA. Patient assessments and online ratings of quality care: a "wake-up call" for providers. Am J Gastroenterol 2013 Nov;108(11):1676-1685. [doi: 10.1038/ajg.2013.112] [Medline: 24192941]
62. Black E, Thompson L, Saliba H, Dawson K, Black NM. An analysis of healthcare providers' online ratings. Inform Prim Care 2009;17(4):249-253. [Medline: 20359403]
63. Kadry B, Chu L, Kadry B, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. J Med Internet Res 2011;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
64. Detz A, López A, Sarkar U. Long-term doctor-patient relationships: patient perspective from online reviews. J Med Internet Res 2013;15(7):e131 [FREE Full text] [doi: 10.2196/jmir.2552] [Medline: 23819959]

65. Alemi F, Torii M, Clementz L, Aron DC. Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments. Qual Manag Health Care 2012;21(1):9-19. [doi: 10.1097/QMH.0b013e3182417fc4] [Medline: 22207014]
66. Lagu T, Hannon N, Rothberg M, Lindenauer P. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
67. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Parental awareness and use of online physician rating sites. Pediatrics 2014 Oct;134(4):e966-e975 [FREE Full text] [doi: 10.1542/peds.2014-0681] [Medline: 25246629]
68. Adams SA. Sourcing the crowd for health services improvement: rhe reflexive patient and "share-your-experience" websites. Soc Sci Med 2011 Apr;72(7):1069-1076. [doi: 10.1016/j.socscimed.2011.02.001] [Medline: 21414701]
69. Mostaghimi A, Crotty B, Landon B. The availability and nature of physician information on the internet. J Gen Intern Med 2010 Nov;25(11):1152-1156 [FREE Full text] [doi: 10.1007/s11606-010-1425-7] [Medline: 20544300]
70. Ellimoottil C, Hart A, Greco K, Quek M, Farooq A. Online reviews of 500 urologists. J Urol 2013 Jun;189(6):2269-2273. [doi: 10.1016/j.juro.2012.12.013] [Medline: 23228385]
71. Emmert M, Maryschok M, Eisenreich S, Schöffski O. Arzt-Bewertungsportale im Internet—Geeignet zur Identifikation guter Arztpraxen? [Websites to assess quality of care—appropriate to identify good physicians?]. Gesundheitswesen 2009 Feb 16;71(04):e18-e27.
72. Strech D, Reimann S. Deutschsprachige Arztbewertungsportale: Der Status quo ihrer Bewertungskriterien, Bewertungstendenzen und Nutzung [German language physician rating sites: the status quo of evaluation criteria, evaluation tendencies and utilization]. Das Gesundheitswesen 2012;74(8-9):502-503.
73. The Nuffield Trust. Rating Providers for Quality: A Policy Worth Pursuing?. London: The Nuffield Trust; 2013 Mar. URL: http://www.nuffieldtrust.org.uk/sites/files/nuffield/publication/130322_ratings_providers_for_quality_summary.pdf [accessed 2014-09-01] [WebCite Cache ID 6aPkLR7vx]
74. Segal J. The role of the Internet in doctor performance rating. Pain Physician 2009;12(3):659-664 [FREE Full text] [Medline: 19461833]

Abbreviations

GP: general practitioner

Edited by G Eysenbach; submitted 14.08.15; peer-reviewed by L Trigg, D Hanauer, S Adams; comments to author 14.09.15; revised version received 16.10.15; accepted 03.11.15; published 08.12.15.

Please cite as:
Patel S, Cain R, Neailey K, Hooberman L
General Practitioners' Concerns About Online Patient Feedback: Findings From a Descriptive Exploratory Qualitative Study in England
J Med Internet Res 2015;17(12):e276
URL: http://www.jmir.org/2015/12/e276/
doi:10.2196/jmir.4989
PMID:26681299

Original Paper

Assessing Pictograph Recognition: A Comparison of Crowdsourcing and Traditional Survey Approaches

Jinqiu Kuang[1], MS; Lauren Argo[1], BFA; Greg Stoddard[2], MS; Bruce E Bray[1], MD; Qing Zeng-Treitler[1,3], PhD

[1]Department of Biomedical Informatics, University of Utah, Salt Lake City, UT, United States

[2]Study Design and Biostatistics Center, University of Utah, Salt Lake City, UT, United States

[3]George E. Wahlen Department of Veterans Affairs Medical Center, Informatics Decision-Enhancement and Analytic Sciences (IDEAS) Center, Salt Lake City, UT, United States

Corresponding Author:
Jinqiu Kuang, MS
Department of Biomedical Informatics
University of Utah
421 Wakara Way, Suite 140
Salt Lake City, UT, 84108
United States
Phone: 1 801 581 4080
Fax: 1 801 581 4297
Email: Jinqiu.kuang@utah.edu

Abstract

Background: Compared to traditional methods of participant recruitment, online crowdsourcing platforms provide a fast and low-cost alternative. Amazon Mechanical Turk (MTurk) is a large and well-known crowdsourcing service. It has developed into the leading platform for crowdsourcing recruitment.

Objective: To explore the application of online crowdsourcing for health informatics research, specifically the testing of medical pictographs.

Methods: A set of pictographs created for cardiovascular hospital discharge instructions was tested for recognition. This set of illustrations (n=486) was first tested through an in-person survey in a hospital setting (n=150) and then using online MTurk participants (n=150). We analyzed these survey results to determine their comparability.

Results: Both the demographics and the pictograph recognition rates of online participants were different from those of the in-person participants. In the multivariable linear regression model comparing the 2 groups, the MTurk group scored significantly higher than the hospital sample after adjusting for potential demographic characteristics (adjusted mean difference 0.18, 95% CI 0.08-0.28, $P<.001$). The adjusted mean ratings were 2.95 (95% CI 2.89-3.02) for the in-person hospital sample and 3.14 (95% CI 3.07-3.20) for the online MTurk sample on a 4-point Likert scale (1=totally incorrect, 4=totally correct).

Conclusions: The findings suggest that crowdsourcing is a viable complement to traditional in-person surveys, but it cannot replace them.

(J Med Internet Res 2015;17(12):e281) doi:10.2196/jmir.4582

KEYWORDS
crowdsourcing; patient discharge summaries; Amazon Mechanical Turk; pictograph recognition; cardiovascular

Introduction

Crowdsourcing has become increasingly popular in the past decade due to its time-saving and cost-effective qualities [1,2]. Crowdsourcing was primarily used by industries to outsource business tasks. More recently, human subject researchers have taken interest in crowdsourcing as a viable alternative approach to traditional methods of participant recruitment. The study domains include, but are not limited to, social behavioral science [3], psychology [4-6], and other health-related sciences [7-14]. Crowdsourcing has also been used to generate annotation gold standards for natural language processing in a variety of technical fields [15-22].

In the biomedical domain, researchers have begun experimenting with crowdsourcing. A recent systematic review of crowdsourcing used for health and medical research argued that utilizing crowdsourcing could improve the quality, cost, and speed of a research project and contributes to novel scientific findings [7]. Leroy et al [8] recruited Amazon Mechanical Turk (MTurk) workers to evaluate the effects of a text simplification algorithm using term familiarity to improve perceived and actual text difficulty. Yu et al [9] also crowdsourced a pictogram evaluation task to MTurk workers and confirmed that crowdsourcing can be used as an effective and inexpensive approach for participatory evaluation of medical pictograms.

MTurk is a large and well-known crowdsourcing service, which has developed into the leading platform for crowdsourcing recruitment [23]. Two primary concerns in the use of MTurk for human subject research are the demographic mix of the study participants and data quality, both of which affect the validity and generalizability of results obtained from MTurk. Demographic composition of the participants is essential to understanding the sampling bias of the study population and the generalizability of results. Early studies indicated that workers reached by MTurk were mostly US based and this population was younger, better educated, and with a higher proportion of females than the general US population [24-26]. A 2010 paper by Eriksson and Simpson [26] reported that a greater proportion of Indian participants were recruited (with 424 respondents from India and 416 from the United States) for their experiment. A demographic survey conducted by Paolacci et al [25] showed that only 47% of workers were from the United States and there were a significant number of Indian participants (34%). However, it is rumored that Amazon stopped approving new international MTurk accounts since early 2013.

Regarding data quality, researchers have attempted to understand the motivations of the MTurk participants and whether it affected data quality [17,27-31]. Given that the median wage of MTurk workers was as low as US $1.38 per hour [32], one may be concerned about the quality of work. However, a number of prior studies have compared the quality of data between MTurk workers and in-laboratory participants in various research studies and suggested that the data collected online were not of poorer quality than data collected from traditional subject pools [1-5,25,33].

Specific to the biomedical domain, concerns associated with the use of crowdsourcing include exclusion of certain populations, such as minors and people with limited or no computer skills [2]. Other concerns are built-in limitations that include (1) sample biases, (2) inability to control participants' environment, and (3) inability to verify participant responses [34]. For researchers who are interested in clinical populations, the prevalence of clinical conditions and clinical characteristics of MTurk workers and the general population may be different. Another issue is that online informed consent documentation is not always read carefully [35]. Despite these concerns, crowdsourcing is a potential alternative to more traditional methods of subject recruitment.

We have been working on improving hospital discharge instructions with automated pictographic illustrations. Hospital discharge instructions are essential to the patients' postdischarge care because these patients and their families are usually responsible for the majority of care after discharge. However, discharge instructions can be difficult for some patients to understand. Previous studies have shown that more than half of patients do not fully understand the content of instructions [36-38]. Illustrations can help enhance patients' comprehension and recall [39-41]. However, not all illustrations lead to better comprehension and recall [39,42,43]. Therefore, high-quality and effective pictographs are needed. We created a set of pictographs and stored them in a system called "Glyph." Glyph automatically illustrates text with analogous pictographs using natural language processing and computer graphics techniques [44]. For Glyph to be effective, we needed to test and ensure that the pictographs it uses are indeed recognizable by patients.

Given crowdsourcing's low cost, high efficiency, and relatively good data quality, we set out to explore its use for clinical pictograph testing and compare it with a traditional recruitment and survey method. We noted that prior clinical informatics studies have not compared the results obtained from traditional subject recruitment and crowdsourcing. In this study, we tested medical pictograph recognition rate using a hospital sample and using an MTurk sample. MTurk was chosen for this study because it is the most well-established and well-studied crowdsourcing service. It also allowed us to closely control participation and measure the quality of participant output.

Methods

As part of the Gylph project, more than 1000 pictographs were developed. Among them, we randomly selected 500 pictographs for testing. These pictographs were first drawn by a professional graphic designer and then reviewed by a team of clinicians and researchers. Field testing with patients/consumers was then performed because the patient/consumer population is very diverse and the developers were inherently biased by their participation in the design. To test pictograph recognition, we designed a set of questionnaires with fill-in-the-blank questions for which study participants were asked to complete discharge instruction sentences based on the pictures shown. A total of 150 different questionnaires were generated, each containing 50 questions, enabling each pictograph to be tested 15 times.

After the University of Utah Institutional Review Board approved the in-person survey study, 100 study participants were recruited from a cafeteria area of the University of Utah Hospital, which is frequented by patients, visitors, and staff. Another 50 study participants were recruited from the Environmental Services Department via convenience sampling. Inclusion criteria for participants included individuals aged 21 years or older and able to speak, read, and write in English. Exclusion criteria included anyone unable to read; having any visual, cognitive, language, or other impairments that would prevent full participation in the study; and anyone who currently or previously worked with discharge instructions in any capacity. Informed consent was obtained from each participant. Once consented, participants were given a randomly selected questionnaire, asked to read the questions, fill in the blanks based on the pictographs, and provide their demographic information including age, gender, race, ethnicity, education level, and first language. Most participants completed the questionnaire in 10-20 minutes. Each participant received a US $10 gift card for participation [45].

In the crowdsourcing study, we tested 486 of the 500 images using MTurk. In all, 12 duplicate pictographs associated with different instructions were eliminated to avoid confusion and another 2 pictographs were inadvertently omitted. Similar to the in-person survey, 150 study participants were recruited from MTurk. We requested 15 human intelligence tasks (HITs) per survey, each survey containing up to 50 images to be identified and 7 demographic questions: gender, age bracket, ethnicity, race, education level, first language, and country of residence. Each survey taker received US $6 to complete the survey. We requested that each survey taker be unique and have a "Masters" qualification, which is defined as "consistently completing HITs of a certain type with a high degree of accuracy across a variety of requesters." Each survey taker was required to answer the questions even if it was a guess. The system prompted the study participants to enter "??" when they could not guess the meaning. The format was "fill-in-the-blank" with a comment box below the sentence (Figure 1).

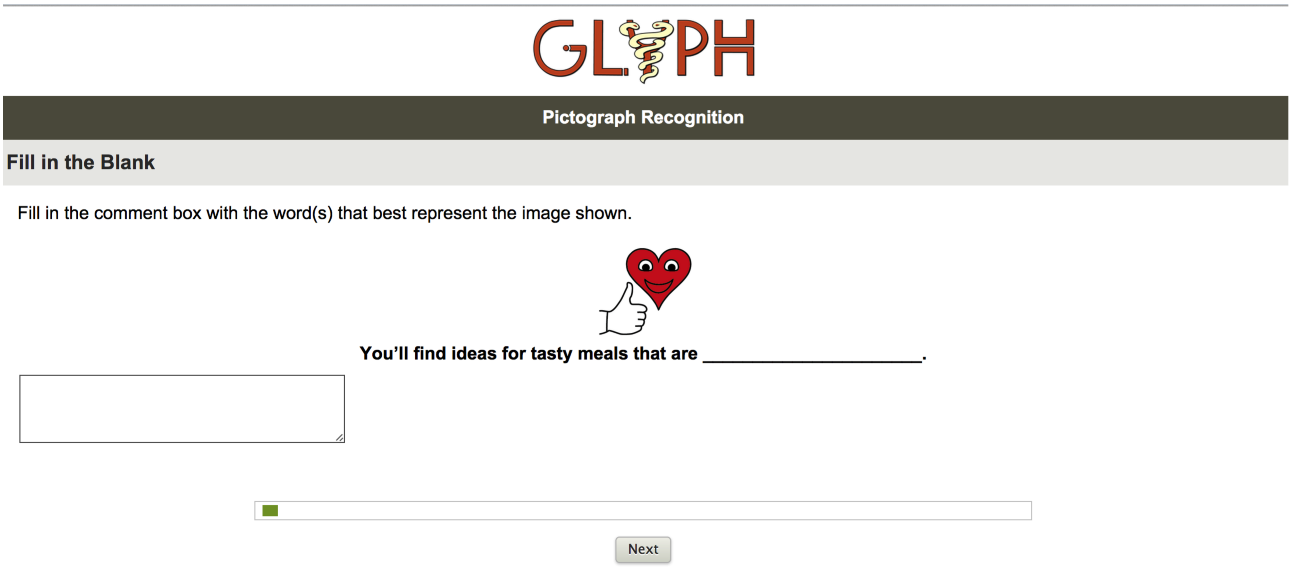
We used SurveyMonkey [46] as the survey creation tool and for analyzing the responses. Verification that the survey takers were all unique was done based on MTurk user IDs.

In the in-person survey study, the questionnaires with handwritten responses were scanned and answers were transcribed into a database. The MTurk answers were collected using SurveyMonkey and later exported to an Excel spreadsheet. Demographic data were coded for statistical analysis. The questionnaire results were evaluated against the phrases used by the original discharge instructions. The following 4-point Likert scale was used: 1=incorrect, 2=mostly incorrect, 3=mostly correct, and 4=correct. Human reviewers first rated 10% of the questionnaires and an interrater agreement was calculated. Disagreements in rating were resolved through consensus. After interrater agreement reached the conventionally acceptable kappa value of .85, individual reviewers independently rated the remaining questionnaires.

Because each pictograph was tested 15 times and each test result was given a rating from 1 to 4, the sum of the ratings for each pictograph ranged from 15 to 60. In this study, we considered a sum of the ratings less than 40 or a mean rating less than 2.67 as "low" or "ineffective," indicating a low recognition rate and need for redesign, whereas a sum of the ratings equal to or greater than 40 points (eg, a mean rating equal to or greater than 2.67) was considered as "high" or "effective," indicating a high recognition rate.

We compared 486 pictographs tested in the crowdsourcing study with their identical counterparts tested in the in-person study. We removed the results in the in-person study that corresponded to the pictographs that were eliminated in the crowdsourcing study due to duplication and omission. We first calculated descriptive statistics for the 2 samples (MTurk and in-person). Mean ratings were then calculated and compared between the in-person hospital sample and online MTurk sample. Afterward, we performed multivariable linear regression analyses to investigate the effects of gender, age, ethnicity, race, education level, and first language on the recognition rates within the 2 populations. Qualitative analyses were then conducted on pictograph characteristics to explore the reasons behind the difference.

Figure 1. Screenshot of a sample question for both groups.



Fill in the Blank

Fill in the comment box with the word(s) that best represent the image shown.

You'll find ideas for tasty meals that are ____________.

Results

Comparing the in-person and crowdsourcing studies, the 2 recruitment groups differed on several demographic characteristics (Table 1). The MTurk sample had more white and less Hispanic participants, were better educated, and had more native English speakers. We did not limit to US workers only, although more than 93.3% (140/150) of the workers were from the United States. There were 10 non-US workers out of 150 (6.7%), all from Asia. There were 18 Asian participants (18/150, 12.0%) in the in-person group. Asian workers had lower recognition rates than white workers did; however, black or African American, American Indian and Alaska Native,

Native Hawaiian and Other Pacific Islander, and other workers scored even lower on average in the in-person study. In our study, the pictograph recognition rate is not a reflection of the participants' effort because we did not observe any sign of lack of effort by a particular population group. In fact, in the in-person survey, those who were less educated and/or who did not speak English as their first language appeared to spend more time completing the questionnaire.

Table 1. Demographics of the in-person and online recruitment groups (N=300).

Demographic characteristics	In-person (n=150)	Online (n=150)	*P*
Gender, n (%)			.23
Male	67 (44.7)	77 (51.7)	
Female	83 (55.3)	72 (48.3)	
Age (years), n (%)			.005
21-29	46 (30.9)	56 (37.3)	
30-39	36 (24.2)	44 (29.3)	
40-49	31 (20.8)	45 (30.0)	
50-59	22 (14.8)	3 (2.0)	
60-69	13 (8.7)	2 (1.3)	
70-79	1 (0.7)	0 (0.0)	
Race, n (%)			<.001
White	86 (57.3)	135 (90.0)	
Asian	18 (12.0)	10 (6.7)	
Other	46 (30.7)	5 (3.3)	
Ethnicity, n (%)			<.001
Hispanic	30 (23.1)	4 (2.7)	
Non-Hispanic	100 (76.9)	144 (97.3)	
Education (grade), n (%)			<.001
≤4	3 (2.0)	0 (0.0)	
5-8	4 (2.7)	0 (0.0)	
9-12	25 (16.7)	15 (10.1)	
>12	118 (78.7)	133 (89.9)	
First language, n (%)			<.001
English	100 (67.1)	139 (92.7)	
Non-English	49 (32.9)	11 (7.3)	

The mean time spent per survey online was 23.9 minutes (95% CI 22.5-25.3), whereas most in-person participants recruited from the hospital cafeteria area completed the questionnaire in 10-20 minutes. This suggests that online workers were not less attentive.

In the multivariable linear regression model comparing the 2 groups (Table 2), online participants scored significantly higher than the in-person participants after adjusting for demographic characteristics. The majority of pictographs scored well in the recognition test: adjusted mean ratings were 2.95 (95% CI 2.89-3.02) for the in-person sample and 3.14 (95% CI 3.07-3.20) for the MTurk sample on the 4-point Likert scale. The adjusted mean difference was 0.18 (95% CI 0.08-0.28, *P*<.001). This suggests that the MTurk responders were better at recognizing the set of pictographs we tested than the hospital sample were and the difference could not be completely explained by the demographic variables we collected.

Table 2. Multivariable linear regression model of mean ratings between the online and in-person groups.

Predictors	Adjusted mean difference (slope)	95% CI	*P*
MTurk	0.18	0.08 to 0.28	<.001
Gender (male)	0.01	−0.07 to 0.10	.75
Age (years)			
21-29	Referent		
30-39	0.14	0.03 to 0.25	.01
40-49	−0.09	−0.20 to 0.02	.11
50-59	0.10	−0.07 to 0.26	.25
60-69	0.01	−0.20 to 0.22	.92
70-79	0.37	−0.33 to 1.08	.29
Race (%)			
White	Referent		
Asian	−0.05	0.24 to 0.14	.59
Other	−0.17	−0.31 to −0.02	.02
Ethnicity (Hispanic)	0.05	2.79 to 3.08	.54
Education (>12th grade/college)	0.15	0.03 to 0.27	.02
First language (non-English)	−0.54	−0.69 to −0.40	<.001

The model presented in Table 2 identified several predictors of recognition rate in addition to study group. For age, compared with the 21-29 year group, the 30-39 year group mean rating was higher by 0.17 (95% CI 0.03-0.25, *P*=.01). Other older age groups were not significantly associated with rating change. Compared with the white participants, the mean rating for Asian participants was not significantly different (0.05, 95% CI −0.24 to 0.14, *P*=.59) and “other” race ratings were 0.17 higher (95% CI −0.31 to −0.02, *P*=.02). Compared with high-school graduates or lower, college graduates’ mean rating was raised by 0.15 (95% CI 0.03-0.27, *P*=.02). Compared with English as first language, mean ratings for those who did not speak English as a first language ratings were lowered by 0.54 (95% CI −0.69 to −0.40, *P*<.001). No significant differences were detected between mean rating and gender (*P*=.75) or ethnicity (*P*=.54).

In the qualitative analysis, we sought to identify general pictographic characteristics that affected recognition by the 2 groups. We examined 3 different categories of pictographs based on recognition ratings. The 3 categories were (1) images that had no variation in mean ratings (n=29), (2) images that scored at least 0.5 points higher in mean ratings with the in-person hospital sample (n=15), and (3) those that scored at least 1 point higher in mean ratings with the online MTurk workers (n=49). Among the 486 pictographs, only 29 had the exact same ratings, although the rating differences were fairly small (<0.5) for the majority of the pictographs. The in-person hospital sample scored higher in 79 images, whereas MTurk workers scored higher in 379 images. Figures 2 and 3 display sample questions and answers with the most similar and the most different scores between the 2 samples.

As part of our analysis, the test pictographs were classified as direct, indirect, and arbitrary according to the representation strategies outlined by Nakamura and Zeng-Treitler [47]. Direct representation explored the visual similarity between a pictograph and its referent, (eg, depicting a thermometer directly). Arbitrary representations were established by social convention (eg, using a red “X” to indicate “no”). Indirect representation explored semantic relations between a pictograph and its referent (eg, using a cactus to represent “dry”). A fourth hybrid category was used for pictographs that contained both indirect and arbitrary elements. Indirect representation was further classified by sematic type. In both samples, the most recognized strategy was direct followed by arbitrary, indirect, and indirect with arbitrary (Table 3). The mean rating within different demographic groups by representation strategy is shown in Table 4. Indirect and indirect with arbitrary strategies were particularly ineffective for older patients, Hispanics, non-Whites, and non-native English speakers. This suggests that the indirect and arbitrary strategies are more culturally dependent.

Table 3. Mean rating by representation strategy.

Representation strategy	n (total=486)	Mean rating (SD)	
		Online	In-person
Direct	165	3.45 (0.90)	3.20 (1.08)
Arbitrary	5	3.35 (1.12)	3.09 (1.24)
Indirect	160	3.18 (1.08)	2.77 (1.22)
Indirect with arbitrary	156	3.04 (1.09)	2.56 (1.22)

Table 4. The mean rating within different demographic groups by representation strategy.

Demographic groups	Mean rating by representation strategy (SD)				Overall mean rating (SD)
	Indirect	Direct	Indirect with arbitrary	Arbitrary	
Gender					
Male	2.95 (1.18)	3.31 (1.00)	2.77 (1.19)	3.21 (1.16)	3.02 (1.15)
Female	2.99 (1.16)	3.34 (1.00)	2.82 (1.18)	3.22 (1.23)	3.05 (1.14)
Age (years)					
21-29	3.04 (1.14)	3.33 (0.99)	2.83 (1.18)	3.21 (1.20)	3.06 (1.13)
30-39	3.04 (1.15)	3.44 (0.92)	2.89 (1.16)	3.36 (1.17)	3.14 (1.10)
40-49	2.90 (1.20)	3.23 (1.08)	2.74 (1.20)	3.10 (1.20)	2.96 (1.18)
50-59	2.80 (1.22)	3.30 (1.03)	2.60 (1.23)	3.40 (1.26)	2.92 (1.20)
60-69	2.83 (1.15)	3.22 (0.98)	2.70 (1.15)	2.80 (1.30)	2.98 (1.10)
70-79	3.22 (1.20)	3.65 (0.69)	2.25 (1.50)	4.00 (0)	3.46 (0.94)
Ethnicity					
Hispanic	2.67 (1.20)	3.14 (1.10)	2.44 (1.19)	3.54 (0.88)	2.76 (1.20)
Non-Hispanic	3.03 (1.16)	3.37 (0.97)	2.86 (1.17)	3.26 (1.17)	3.09 (1.12)
Race					
Nonwhite	2.52 (1.21)	3.01 (1.16)	2.34 (1.19)	2.93 (1.25)	2.63 (1.22)
White	3.13 (1.12)	3.44 (0.92)	2.95 (1.14)	3.28 (1.18)	3.18 (1.08)
Education					
≤12th grade	2.78 (1.17)	3.14 (1.10)	2.48 (1.19)	2.67 (1.34)	2.81 (1.19)
>12th grade	3.01 (1.17)	3.36 (0.98)	2.85 (1.18)	3.33 (1.13)	3.08 (1.13)
First language					
English	3.12 (1.12)	3.44 (0.92)	2.95 (1.14)	3.32 (1.14)	3.18 (1.08)
Non-English/other	2.41 (1.21)	2.87 (1.19)	2.22 (1.18)	2.85 (1.26)	2.50 (1.23)

There were 29 images that received the same recognition scores in both samples. These were generally high-scoring images that represented common objects, activities, behaviors, and common disorders. It should be mentioned there were 3 low scoring images in this category that were not recognizable by either group. This group of pictographs largely represented simple ideas and common behaviors with the use of large fields of open space (Figure 4). Many of these used the direct representation strategy; however, the use of arbitrary symbols was successful in many cases as well.

Although the majority of pictographs scored much higher on average with the online group, there were 15 images that scored at least 0.5 points higher with the in-person group. These images tended to have more contrast using color and did not represent overly complex or abstract ideas. With an overall mean rating of 2.95 (SD 0.65) in both groups, these were recognizable in general (Figure 5).

The final and largest category contained the 49 pictographs that scored at least 1 point or higher by the online group. These images attempted to communicate more complex or abstract concepts than the other 2 categories. The mean rating for the pictographs in this category was 2.77 (SD 0.59) for the in-person sample and 3.31 (SD 0.59) for the online sample. Almost every pictograph in this category used the indirect representation strategy. This category also had many pictographs that contained fine detail and the use of color was not as prevalent as in the other categories (Figure 6). One example is “difficulty sleeping”; although “sleeping” can be easily illustrated, “difficulty” is an abstract and challenging concept to visualize.

These results suggest that the online sample was better at recognizing complex and/or abstract ideas communicated through images. It might be that the MTurks were able to improve their recognition rating by zooming into the screen to see more finely detailed images or they were more familiar with visual icons. Overall, the most efficient way to communicate visually to a diverse audience was the direct representation strategy, employing simple concepts with color and heavy contrast.

Discussion

This study is the first effort to compare the results of conventional and crowdsourcing recruitment in the health informatics domain. Our crowdsourcing (MTurk) sample had different demographic characteristics from the conventional (hospital patient, visitor, and staff) sample. After adjusting for demographic variables, the crowdsourcing (online) sample scored higher on the pictograph recognition tasks than the conventional (in-person) sample (P<.001). This suggests that we cannot simply replace conventional recruitment with crowdsourcing recruitment.

At the same time, the crowdsourcing recruitment was much cheaper and quicker. The data quality was also relatively high. We found no missing data and no transcription was needed. For many pictographs, the differences in the average recognition scores were not dramatic. This suggests that online crowdsourcing is a viable approach for preliminary pictograph evaluation.

Crowdsourcing services—particularly MTurk—have made it easy for scientists to recruit research participants. However, we should not overlook the crucial differences between crowdsourcing and traditional recruitment methods. Not all the tasks that are performed in an in-person setting are suitable for crowdsourcing online. Current general crowdsourcing tools are not specifically tailored for biomedical informatics research. From a human subject researcher’s standpoint, representativeness of a sample is critical. However, tools such as MTurk or SurveyMonkey provide limited capabilities for researchers to sample subjects that mimic the target population of a specific research project. Along the same line, it remains to be explored how crowdsourcing can be incorporated into longitudinal and/or intervention studies.

Creating high-quality and effective pictographs is our goal. To achieve this goal, an iterative process of design and testing was carried out. User testing is intended to identify pictographs that are confusing, allowing those pictographs to be redesigned and retested. In other words, the purpose of the pictograph recognition test is to assess the quality of the pictographs rather than to assess the knowledge and skill of the users. As such, the quality of pictographs being tested will vary and the “wrong” answers are as valuable to us as the “correct” answers.

Figure 2. Sample questions with the same or the most different scores from the online and in-person samples.

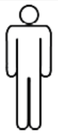
Image	Sample question	Sum of rating		Correct answer
		Online	In-person	
	Maintain a ________________. Get help to lose any extra pounds.	22	42	Healthy weight
	Beta-blocker-- ________________, slows heart rate, keeps rhythm regular.	49	23	Lowers blood pressure
	Mild side effects include ________________.	60	60	Vomiting
	After heart surgery, you may ________________ for several months especially in the arm or leg they took blood vessels from.	26	26	Retain fluid

Figure 3. Answers (n=15) to the sample questions from the online and in-person samples.

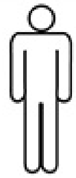
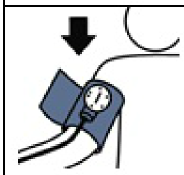
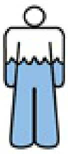
Image	Correct answer	Answers	
		Online	In-person
	Healthy weight	Upright posture (n=2) Healthy lifestyle Straight posture (n=4) Good posture (n=2) Standing position (n=2) Healthy diet Maintain a good health Upright composure Balance	Good weight (n=3) Proper body weight Normal weight Man? Normal diet Healthy diet Healthy weight Good posture Suitable weight Body (n=2) Weight (n=2)
	Lowers blood pressure	Regulates blood pressure Pressure Blood pressure medications Lowers blood pressure Checks blood pressure Blood pressure (n=2) Decreases blood pressure Low blood pressure (n=2) Helps to Cuff Decreases blood pressure Which lower blood pressure Blood pressure suppressant	Blood pressure (n=5) Decreases bp Blood flow ? (n=4) Medication Regular Blood Machine
	Vomiting	Vomiting (n=15)	Vomiting (n=10) Vomit (n=5)
	Retain fluid	Be bloated Experience water retention Feel cold (n=6) Feel cold in the lower body Feel coldness Retain water (n=2) Chills Feel numbness Feel weak	Feel pain Discoloration Have fluid Feel light headed Feel cold Have swellies Retain water (n=2) Have cold limbs Feel numbness (n=2) Bleed Be sore ? Feel week (weak)

Figure 4. Pictographs that scored the same for both the online and in-person groups.

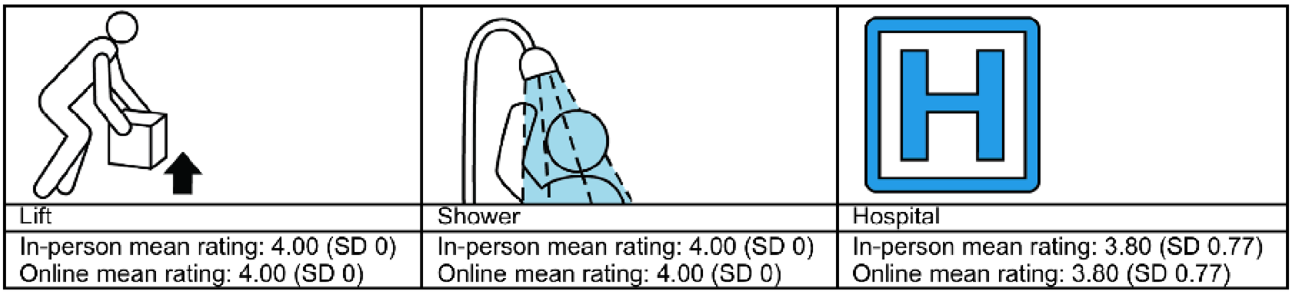


Figure 5. Pictographs that scored higher with the in-person group.

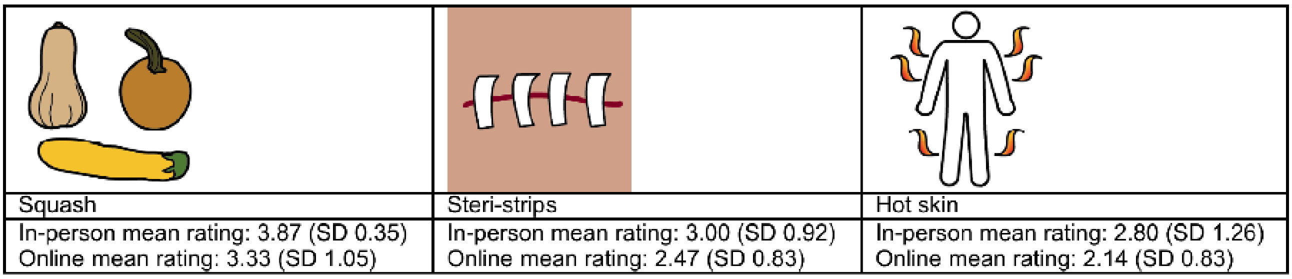
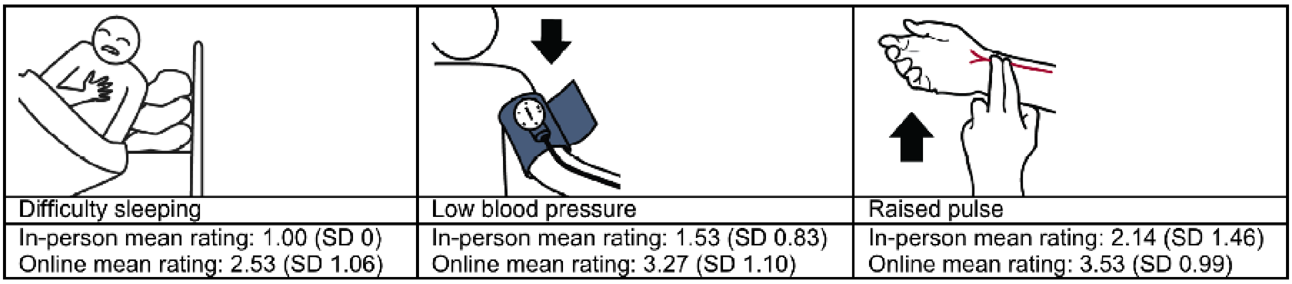


Figure 6. Pictographs that scored higher with the online group.



This study has some limitations. We focused on a single task (pictograph recognition) and a single crowdsourcing service (MTurk). Our conventional sample was recruited from a hospital where our target audience for the pictograph-enhanced instructions receives care. Arguably, a sample recruited from a different location in the United States and a different type of health care facility will have different characteristics and different recognition rates.

In future studies, especially in informatics studies that target patients, we plan to further explore the use of crowdsourcing services. For instance, one of our ongoing projects aims at reducing the disparity in health communication through pictographs. Crowdsourcing is a method that could potentially help us recruit participants from more diverse groups and develop pictographs that are more widely recognizable.

We tested the recognition of health care-related pictographs through crowdsourcing and conventional in-person survey. The self-reported demographics of our online MTurk workers indicated they were younger and more educated than the conventional in-person survey sample. The majority were white and English was their first language. Despite the demographic differences between the 2 study groups, predictors of successful pictograph recognition remain the same: white, college educated, and native English language speaking.

Crowdsourcing has some distinct advantages: it is time-saving, low cost, and less labor intensive (for the researchers). However, our analyses indicated that after adjusting for demographic characteristics, the average pictograph recognition rating of online MTurk and in-person hospital survey participants was significantly different. Therefore, the crowdsourcing approach cannot simply replace conventional survey methods, although it could be used for preliminary studies and quick feedback.

Acknowledgments

This work was supported by NIH grants R01 LM07222 and 5G08LM11546. We thank all the Turkers and individuals who participated in this study.

Conflicts of Interest

None declared.

References

1. Saunders DR, Bex PJ, Woods RL. Crowdsourcing a normative natural language dataset: A comparison of Amazon Mechanical Turk and in-lab data collection. J Med Internet Res 2013;15(5):e100 [FREE Full text] [doi: 10.2196/jmir.2620] [Medline: 23689038]
2. Azzam T, Jacobson MR. Finding a comparison group: Is online crowdsourcing a viable option? Am J Eval 2013 Jun 18;34(3):372-384. [doi: 10.1177/1098214013490223]
3. Behrend TS, Sharek DJ, Meade AW, Wiebe EN. The viability of crowdsourcing for survey research. Behav Res Methods 2011 Sep;43(3):800-813. [doi: 10.3758/s13428-011-0081-0] [Medline: 21437749]
4. Krantz JH, Dalal R. Validity of Web-based psychological research. In: Birnbaum MH, editor. Psychological Experiments on the Internet. San Diego, CA: Academic Press; 2000:35-60.
5. Gosling SD, Vazire S, Srivastava S, John OP. Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. Am Psychol 2004;59(2):93-104. [doi: 10.1037/0003-066X.59.2.93] [Medline: 14992636]
6. Shapiro DN, Chandler J, Mueller PA. Using Mechanical Turk to study clinical populations. Clinical Psychological Science 2013:213-220.
7. Ranard BL, Ha YP, Meisel ZF, Asch DA, Hill SS, Becker LB, et al. Crowdsourcing--harnessing the masses to advance health and medicine, a systematic review. J Gen Intern Med 2014 Jan;29(1):187-203 [FREE Full text] [doi: 10.1007/s11606-013-2536-8] [Medline: 23843021]
8. Leroy G, Endicott JE, Kauchak D, Mouradi O, Just M. User evaluation of the effects of a text simplification algorithm using term familiarity on perception, understanding, learning, and information retention. J Med Internet Res 2013;15(7):e144 [FREE Full text] [doi: 10.2196/jmir.2569] [Medline: 23903235]
9. Yu B, Willis M, Sun P, Wang J. Crowdsourcing participatory evaluation of medical pictograms using Amazon Mechanical Turk. J Med Internet Res 2013;15(6):e108 [FREE Full text] [doi: 10.2196/jmir.2513] [Medline: 23732572]
10. Allam A, Kostova Z, Nakamoto K, Schulz PJ. The effect of social support features and gamification on a Web-based intervention for rheumatoid arthritis patients: Randomized controlled trial. J Med Internet Res 2015;17(1):e14 [FREE Full text] [doi: 10.2196/jmir.3510] [Medline: 25574939]
11. Brady CJ, Villanti AC, Pearson JL, Kirchner TR, Gupta OP, Shah CP. Rapid grading of fundus photographs for diabetic retinopathy using crowdsourcing. J Med Internet Res 2014;16(10):e233 [FREE Full text] [doi: 10.2196/jmir.3807] [Medline: 25356929]
12. Cui L, Carter R, Zhang GQ. Evaluation of a novel conjunctive exploratory navigation interface for consumer health information: A crowdsourced comparative study. J Med Internet Res 2014;16(2):e45 [FREE Full text] [doi: 10.2196/jmir.3111] [Medline: 24513593]
13. Helander E, Kaipainen K, Korhonen I, Wansink B. Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: Retrospective cohort study. J Med Internet Res 2014;16(4):e109 [FREE Full text] [doi: 10.2196/jmir.3084] [Medline: 24735567]
14. Turner AM, Kirchhoff K, Capurro D. Using crowdsourcing technology for testing multilingual public health promotion materials. J Med Internet Res 2012;14(3):e79 [FREE Full text] [doi: 10.2196/jmir.2063] [Medline: 22664384]
15. Finin T, Murnane W, Karandikar A, Keller N, Martineau J, Dredze M. Annotating named entities in Twitter data with crowdsourcing. Stroudsburg, PA: Association for Computational Linguistics; 2010 Presented at: Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk; Jun 6, 2010; Los Angeles, CA p. 80-88.
16. Lawson N, Eustice K, Perkowitz M, Yetisgen-Yildiz M. Annotating large email datasets for named entity recognition with Mechanical Turk. Stroudsburg, PA: Association for Computational Linguistics; 2010 Presented at: Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk; Jun 6, 2010; Los Angeles, CA p. 71-79.
17. Snow R, O'Connor B, Jurafsky D, Ng A. Cheap and fast—But is it good?: Evaluating non-expert annotations for natural language tasks. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing. 2008 Presented at: Conference on Empirical Methods in Natural Language Processing; Oct 25-27, 2008; Honolulu, HI p. 254-263.

18. Good BM, Loguercio S, Griffith OL, Nanis M, Wu C, Su AI. The cure: Design and evaluation of a crowdsourcing game for gene selection for breast cancer survival prediction. JMIR Serious Games 2014;2(2):e7 [FREE Full text] [doi: 10.2196/games.3350] [Medline: 25654473]
19. Good BM, Nanis M, Wu C, Su AI. Microtask crowdsourcing for disease mention annotation in PubMed abstracts. Pac Symp Biocomput 2015:282-293 [FREE Full text] [Medline: 25592589]
20. Good BM, Su AI. Crowdsourcing for bioinformatics. Bioinformatics 2013 Aug 15;29(16):1925-1933 [FREE Full text] [doi: 10.1093/bioinformatics/btt333] [Medline: 23782614]
21. Yetisgen-Yildiz M, Solti I, Xia F. Using Amazon's Mechanical Turk for annotating medical named entities. AMIA Annu Symp Proc 2010;2010:1316 [FREE Full text] [Medline: 21785667]
22. Zhai H, Lingren T, Deleger L, Li Q, Kaiser M, Stoutenborough L, et al. Web 2.0-based crowdsourcing for high-quality gold standard development in clinical natural language processing. J Med Internet Res 2013;15(4):e73 [FREE Full text] [doi: 10.2196/jmir.2426] [Medline: 23548263]
23. Amazon Mechanical Turk. Services URL: https://www.mturk.com/mturk/welcome [accessed 2015-09-28] [WebCite Cache ID 6btHG43Sd]
24. Ross J, Irani L, Silberman M, Zaldivar A, Tomlinson B. Who are the crowdworkers?: Shifting demographics in Mechanical Turk. In: CHI '10 Extended Abstracts on Human Factors in Computing Systems. 2010 Presented at: CHI EA '10; Apr 10-15, 2010; Atlanta, GA.
25. Paolacci G, Chandler J, Ipeirotis PG. Running experiments on Amazon Mechanical Turk. Judgment and Decision Making 2010;5(5):411-419.
26. Eriksson K, Simpson B. Emotional reactions to losing explain gender differences in entering a risky lottery. Judgm Decis Mak 2010;5(3):159-163.
27. Shaw A. CrowdFlower Blog. 2010 Aug 05. For love or for money? A list experiment on the motivations behind crowdsourcing work URL: https://www.crowdflower.com/blog/2010/08/for-love-or-for-money-a-list-experiment-on-the-motivations-behind-crowdsourcing-work [accessed 2015-12-11] [WebCite Cache ID 6dhjhsrTX]
28. Chandler D, Kapelner A. Breaking monotony with meaning: Motivation in crowdsourcing markets. J Econ Behav Organ 2013 Jun;90:123-133. [doi: 10.1016/j.jebo.2013.03.003]
29. Callison-Burch C. Fast, cheap, and creative: Evaluating translation quality using Amazon Mechanical Turk. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing. 2009 Presented at: Conference on Empirical Methods in Natural Language Processing; Aug 6-7, 2009; Singapore p. 286-295.
30. Kittur A, Chi EH, Suh B. Crowdsourcing user studies with Mechanical Turk. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2008 Presented at: CHI '08 SIGCHI Conference on Human Factors in Computing Systems; Apr 05-10, 2008; Florence, Italy.
31. Ipeirotis PG. Analyzing the Amazon Mechanical Turk marketplace. XRDS 2010 Dec 01;17(2):16-21. [doi: 10.1145/1869086.1869094]
32. Horton JJ, Chilton LB. The labor economics of paid crowdsourcing. In: Proceedings of the 11th ACM conference on Electronic commerce. 2010 Presented at: EC '10 11th ACM Conference on Electronic Commerce; Jun 9-11, 2010; Cambridge, MA p. 209-218. [doi: 10.1145/1807342.1807376]
33. Sprouse J. A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. Behav Res Methods 2011 Mar;43(1):155-167 [FREE Full text] [doi: 10.3758/s13428-010-0039-7] [Medline: 21287108]
34. Kraut R, Olson J, Banaji M, Bruckman A, Cohen J, Couper M. Psychological research online: Report of Board of Scientific Affairs' Advisory Group on the Conduct of Research on the Internet. Am Psychol 2004;59(2):105-117. [doi: 10.1037/0003-066X.59.2.105] [Medline: 14992637]
35. Stanton JM, Rogelberg SG. Using internet/intranet web pages to collect organizational research data. Organ Res Meth 2001 Jul 01;4(3):200-217. [doi: 10.1177/109442810143002]
36. Jolly BT, Scott JL, Sanford SM. Simplification of emergency department discharge instructions improves patient comprehension. Ann Emerg Med 1995 Oct;26(4):443-446. [Medline: 7574126]
37. Austin PE, Matlack R, Dunn KA, Kesler C, Brown CK. Discharge instructions: Do illustrations help our patients understand them? Ann Emerg Med 1995 Mar;25(3):317-320. [Medline: 7532382]
38. Engel K, Buckley BA, Forth VE, McCarthy DM, Ellison EP, Schmidt MJ, et al. Patient understanding of emergency department discharge instructions: Where are knowledge deficits greatest? Acad Emerg Med 2012 Sep;19(9):E1035-E1044 [FREE Full text] [doi: 10.1111/j.1553-2712.2012.01425.x] [Medline: 22978730]
39. Mayer RE, Gallini JK. When is an illustration worth ten thousand words? J Educ Psychol 1990;82(4):715-726. [doi: 10.1037/0022-0663.82.4.715]
40. Kools M, van de Wiel MW, Ruiter RA, Kok G. Pictures and text in instructions for medical devices: Effects on recall and actual performance. Patient Educ Couns 2006 Dec;64(1-3):104-111. [doi: 10.1016/j.pec.2005.12.003] [Medline: 16472960]
41. Morrow D, Hier CM, Menard WE, Leirer VO. Icons improve older and younger adults' comprehension of medication information. J Gerontol B Psychol Sci Soc Sci 1998 Jul;53(4):P240-P254. [Medline: 9679516]

42. Hwang SW, Tram CQ, Knarr N. The effect of illustrations on patient comprehension of medication instruction labels. BMC Fam Pract 2005 Jun 16;6(1):26 [FREE Full text] [doi: 10.1186/1471-2296-6-26] [Medline: 15960849]
43. Lajoie SP. Extending the scaffolding metaphor. Instr Sci 2005 Nov;33(5-6):541-557. [doi: 10.1007/s11251-005-1279-2]
44. Bui D, Nakamura C, Bray BE, Zeng-Treitler Q. Automated illustration of patients instructions. AMIA Annu Symp Proc 2012;2012:1158-1167 [FREE Full text] [Medline: 23304392]
45. Perri S, Argo L, Kuang J, Bui D, Hill B, Bray EB, et al. A picture's meaning: The design and evaluation of pictographs illustrating patient discharge instructions. J Commun Healthc 2015:e1 (forthcoming).
46. SurveyMonkey. URL: http://www.surveymonkey.com [WebCite Cache ID 6cKfrADBt]
47. Nakamura C, Zeng-Treitler Q. A taxonomy of representation strategies in iconic communication. Int J Hum Comput Stud 2012 Aug 1;70(8):535-551 [FREE Full text] [doi: 10.1016/j.ijhcs.2012.02.009] [Medline: 22754274]

Abbreviations

HIT: human intelligence task
MTurk: Amazon Mechanical Turk

Edited by G Eysenbach; submitted 27.04.15; peer-reviewed by G Leroy, C Smith; comments to author 19.08.15; revised version received 16.10.15; accepted 05.11.15; published 17.12.15.

Please cite as:
Kuang J, Argo L, Stoddard G, Bray BE, Zeng-Treitler Q
Assessing Pictograph Recognition: A Comparison of Crowdsourcing and Traditional Survey Approaches
J Med Internet Res 2015;17(12):e281
URL: http://www.jmir.org/2015/12/e281/
doi:10.2196/jmir.4582
PMID:26678085

Original Paper

Establishing a Link Between Prescription Drug Abuse and Illicit Online Pharmacies: Analysis of Twitter Data

Takeo Katsuki[1*], PhD; Tim Ken Mackey[2,3,4*], MAS, PhD; Raphael Cuomo[4,5], MPH

[1]Kavli Institute for Brain and Mind, University of California, San Diego, La Jolla, CA, United States
[2]Department of Anesthesiology, UC San Diego - School of Medicine, La Jolla, CA, United States
[3]Division of Global Public Health, UC San Diego - School of Medicine, La Jolla, CA, United States
[4]Global Health Policy Institute, La Jolla, CA, United States
[5]Joint Doctoral Program in Global Public Health, UC San Diego - San Diego State University, La Jolla, CA, United States
[*]these authors contributed equally

Corresponding Author:
Tim Ken Mackey, MAS, PhD
Global Health Policy Institute
8950 Villa La Jolla Drive
A204
La Jolla, CA, 92037
United States
Phone: 1 858 264 0163
Fax: 1 858 534 9168
Email: tmackey@ucsd.edu

Abstract

Background: Youth and adolescent non-medical use of prescription medications (NUPM) has become a national epidemic. However, little is known about the association between promotion of NUPM behavior and access via the popular social media microblogging site, Twitter, which is currently used by a third of all teens.

Objective: In order to better assess NUPM behavior online, this study conducts surveillance and analysis of Twitter data to characterize the frequency of NUPM-related tweets and also identifies illegal access to drugs of abuse via online pharmacies.

Methods: Tweets were collected over a 2-week period from April 1-14, 2015, by applying NUPM keyword filters for both generic/chemical and street names associated with drugs of abuse using the Twitter public streaming application programming interface. Tweets were then analyzed for relevance to NUPM and whether they promoted illegal online access to prescription drugs using a protocol of content coding and supervised machine learning.

Results: A total of 2,417,662 tweets were collected and analyzed for this study. Tweets filtered for generic drugs names comprised 232,108 tweets, including 22,174 unique associated uniform resource locators (URLs), and 2,185,554 tweets (376,304 unique URLs) filtered for street names. Applying an iterative process of manual content coding and supervised machine learning, 81.72% of the generic and 12.28% of the street NUPM datasets were predicted as having content relevant to NUPM respectively. By examining hyperlinks associated with NUPM relevant content for the generic Twitter dataset, we discovered that 75.72% of the tweets with URLs included a hyperlink to an online marketing affiliate that directly linked to an illicit online pharmacy advertising the sale of Valium without a prescription.

Conclusions: This study examined the association between Twitter content, NUPM behavior promotion, and online access to drugs using a broad set of prescription drug keywords. Initial results are concerning, as our study found over 45,000 tweets that directly promoted NUPM by providing a URL that actively marketed the illegal online sale of prescription drugs of abuse. Additional research is needed to further establish the link between Twitter content and NUPM, as well as to help inform future technology-based tools, online health promotion activities, and public policy to combat NUPM online.

(J Med Internet Res 2015;17(12):e280) doi:10.2196/jmir.5144

KEYWORDS

social media; surveillance; prescription drug abuse; twitter; eHealth; illicit Internet pharmacies; cyberpharmacies; infodemiology; infoveillance

Introduction

Prescription drug abuse among youth and adolescents is a recognized national public health crisis [1]. Current data and behavioral trends on "non-medical use of prescription medication" (NUPM) are largely derived from nationally representative anonymous self-administered surveys that ask American high school students to self-report recent and past drug abuse behavior [2,3]. However, the rapid increase in Internet use, social media engagement, and near universal access to mobile devices among teens (aged 13-17) allows for augmentation of traditional NUPM survey data with other digital sources of information that are readily available for "big data" analysis and that can be used for health surveillance and prevention [4,5]. Specifically, Internet users are increasingly "self-reporting" their behavior on a variety of health subjects outside of structured surveys via multiple online social networking channels, including platforms such as Twitter, Facebook, Instagram, blogs, and other social sharing sites. In order to leverage this secondary source of information that can contribute to a better understanding of NUPM, this study identifies, characterizes, and describes prescription drug abuse trends and behavior via the popular microblogging platform, Twitter, which has already been associated with risky health behavior and is heavily populated by youth and adolescents [1,6,7].

Twitter currently commands some 316 million active monthly users and, though not the predominant social media site among teens, is used by an estimated one third of this age demographic and by 23% of all online adults, thereby serving as an important social and communication information-based research tool [4,8]. Additionally, compared to other social network platforms, Twitter provides one of the most versatile public application programming interfaces (APIs), allowing users to access large-scale real-time and historical communication data, though certain limitations in collecting such data exists (discussed below).

Hence, this study seeks to leverage the ability to access, construct, and analyze large conversational datasets from Twitter in order to assess how NUPM is being promoted in this environment of Internet users, similar to previous studies using Twitter to address other important public health issues, including drug safety [9-17]. The study also expands on prior studies assessing the association between NUPM and Twitter by examining whether the content of user-generated tweets directly enables NUPM access to prescription medications from illicit online pharmacies [7,10,18,19]. Illegal marketing and sales of prescription drugs by online pharmacies is an important public health and patient safety issue that the World Health Organization, US Food and Drug Administration, US Drug Enforcement Agency, and other stakeholders recognize as needing to be addressed [1,6,7,19].

Methods

The methods for this study consist of two distinct phases: data collection (Phase 1) and data coding, analysis, and visualization (Phase 2). We describe each of these phases in detail below.

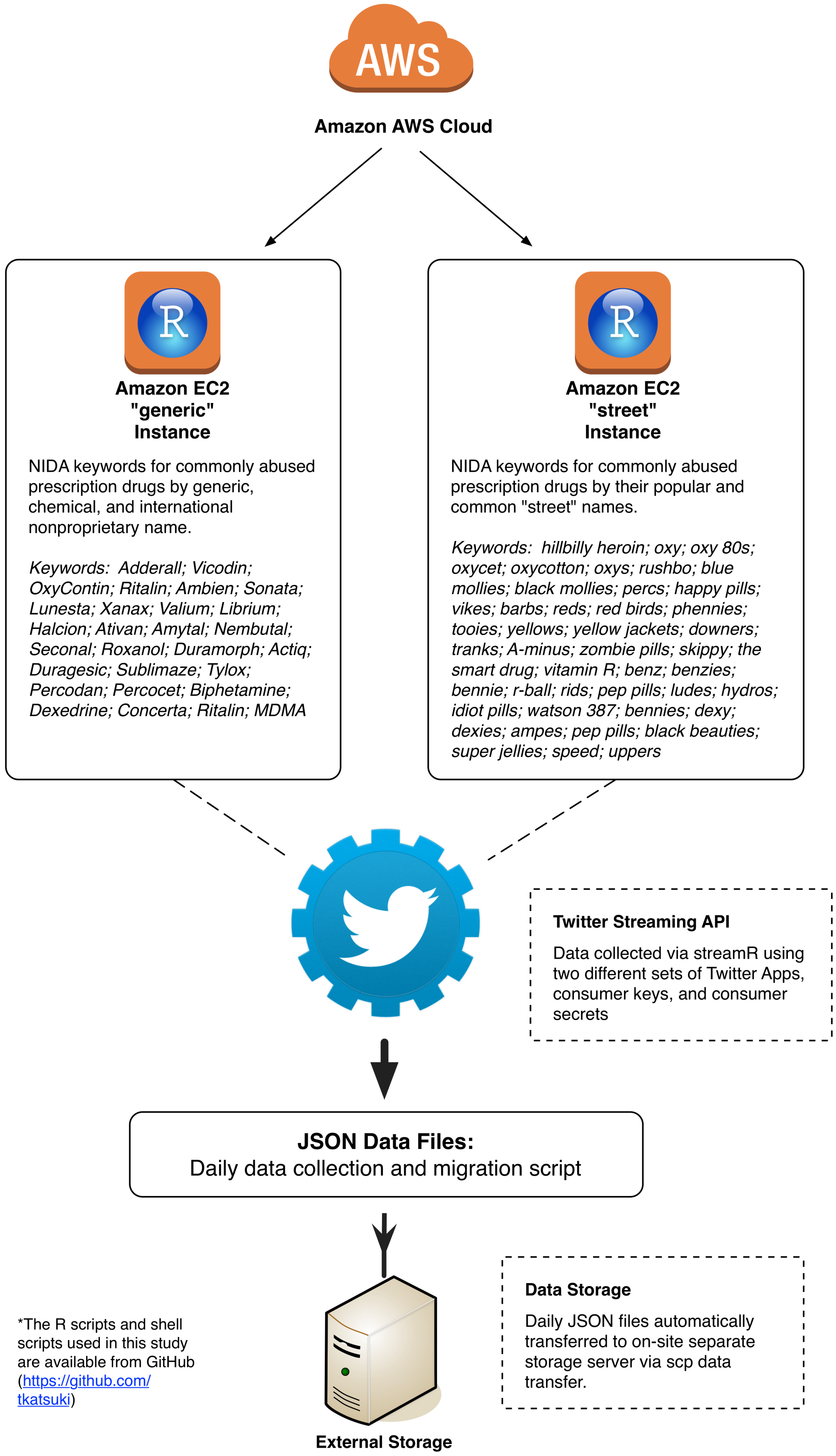
Phase 1: Data Collection

Phase 1 of the study first identified prescription drugs commonly abused by youth and adolescents using information available from the National Institute on Drug Abuse and developed keywords as filters that were then applied to the collection of Twitter data (see Figure 1 for keywords used and visual depiction of the data collection strategy) [20]. We used the identified drug's generic/chemical/international nonproprietary name (eg, oxycodone) and brand name (eg, OxyContin, Percocet) in one set of data collected (ie, Generic Names), and the common "street" or "slang" names (eg, oxy, oxycotton) of drugs in another set of data collection (ie, Street Names) in order to optimize conversational data capture associated with NUPM promotion and behavior [21,22]. Data were collected from the public Twitter Streaming API, and we applied the identified keywords/filters as endpoints in the data capture. This provided us with multiple raw JavaScript Object Notation (JSON) datasets of Twitter feeds and associated metadata for further analysis.

The study conducted an analysis of a 2-week subset of data collected and analyzed using this process from April 1-14, 2015 (ie, Study Data). The two separate datasets of tweets (one filtered for a drug's generic name and a second for street names) were collected from the Twitter Streaming API using streamR package in R (CRAN), which was deployed on cloud-based computing services offered by Amazon Web Services (AWS) via Amazon EC2 t2.micro instances. In accessing the Twitter Streaming API, we used two different sets of Twitter apps, Consumer Keys (API Keys) and Consumer Secrets (API Secrets), in order to maximize data capture and lower the chance of hitting the Twitter Streaming API cap. The 2-week subset of data is part of a larger Twitter NUPM data mining project that has collected 3 months' worth of data and that is undergoing separate analysis. AWS services were chosen due to their relative low cost (discussed below) and primarily for their stability in collecting, transferring, and storing data generated for this project. Specifically, the reliability of AWS (guaranteed availability of 99.95% for external connectivity) ensures contiguity of data when using multiple instances to collect data from the Twitter Streaming API.

R for streaming Twitter data was run on an RStudio Server preconfigured on an Amazon Machine Image (ami-45c72a01) originally developed and made freely accessible to the public (Louis Aslett's RStudio Server Amazon Machine Image website). Streaming was scheduled to iteratively initiate and end every 24 hours, generating daily JSON files that included Twitter data filtered for prescription drug abuse keywords. In the event that streaming was interrupted for any reason, the script was written to automatically prompt the restart of the streaming collection process. The daily files were automatically transferred to a separate file storage server via SCP data transfer, and original files on the AWS server were deleted via SSH if the transfer was successful. Data analysis was performed on a local machine (Dell Precision T5810, 64GB memory, 4 CPU cores) or on an Amazon EC2 m4.4xlarge instance (64GB memory, 16 CPU cores). The R scripts and shell scripts used in this study are available from the first author's GitHub repository (TK).

Figure 1. Data collection strategy.



NIDA keywords for commonly abused prescription drugs by generic, chemical, and international nonproprietary name.

Keywords: Adderall; Vicodin; OxyContin; Ritalin; Ambien; Sonata; Lunesta; Xanax; Valium; Librium; Halcion; Ativan; Amytal; Nembutal; Seconal; Roxanol; Duramorph; Actiq; Duragesic; Sublimaze; Tylox; Percodan; Percocet; Biphetamine; Dexedrine; Concerta; Ritalin; MDMA

NIDA keywords for commonly abused prescription drugs by their popular and common "street" names.

Keywords: hillbilly heroin; oxy; oxy 80s; oxycet; oxycotton; oxys; rushbo; blue mollies; black mollies; percs; happy pills; vikes; barbs; reds; red birds; phennies; tooies; yellows; yellow jackets; downers; tranks; A-minus; zombie pills; skippy; the smart drug; vitamin R; benz; benzies; bennie; r-ball; rids; pep pills; ludes; hydros; idiot pills; watson 387; bennies; dexy; dexies; ampes; pep pills; black beauties; super jellies; speed; uppers

Data collected via streamR using two different sets of Twitter Apps, consumer keys, and consumer secrets

Daily JSON files automatically transferred to on-site separate storage server via scp data transfer.

*The R scripts and shell scripts used in this study are available from GitHub (https://github.com/tkatsuki)

Phase 2: Data Coding, Analysis, and Visualization

Phase 2 of this study involved analyzing data for characteristics of interest by conducting data content coding using a supervised machine learning protocol. The process was first carried out by the second and third authors who acted as human coders and independently reviewed and coded a subset of 1000 randomly selected tweets from each instance. The second author, with expertise and training in substance abuse behavior, trained the third author for content coding. These randomly sampled tweets (including the textual content and select metadata) were reviewed and coded for the following characteristics: (1) relevance to NUPM behavior (ie, reviewing tweets and assessing if they actually discussed NUPM behavior and/or promotion) and (2) assessing NUPM characterization (positive or negative promotion/attitudes) (see Table 1 for details on content coding).

Table 1. Content analysis categories.

Relevant vs non-relevant	Favorable, non-favorable, neutral content analysis	Illicit online source information
Relevant: Contained content discussing NUPM behavior, attitudes, information about buying online, reporting health effects	Favorable promotion: Emphasizing benefits and/or minimizing risks regarding NUPM and generally promoting NUPM lifestyle/behavior	Online access: Providing a URL/hyperlink to “buy” or “online purchase” of prescription drugs
Non-relevant: Topics not associated with NUPM (eg, sports, consumer goods, news reports, music, lawful use of drugs in clinical settings) and tweets without sufficient content to code	Non-favorable promotion: Providing information on risk, side effects, or information on addiction treatment	Risk characteristics of online pharmacy linked to content: Online pharmacy identified as “unapproved” or “rogue” on LegitScript site[a]

[a] LegitScript: “rogue” is categorized as a website that appears to be intentionally or knowingly violating applicable laws or regulations; “unapproved” is categorized as verified as lacking compliance with LegitScript standards or other applicable laws and regulations.

This subset of human coded tweets was then used to train machine classifiers for “relevance” and “favorability” by a Support Vector Machine (SVM) algorithm in R that was then applied to the full dataset of collected tweets. Accuracy of the models was tested with 10-fold cross-validations with 3 repeats using the caret package [23]. In order to create a feature vector representation of each document (tweet), a corpus of the subsampled tweets was generated using the tm package in R [24]. This process involved data cleaning by transforming all texts to lower case, removing uniform resource locators (URLs), numbers, punctuations, and stop words (English language plus “re” and “rt”), as well as word grouping (n-grams) followed by generation of a Term-Document Matrix.

Qualitative analysis of tweets was then conducted by analyzing the source JSON Twitter data in streamR package that parses JSON files and transforms them in R data frames. Additional analysis was conducted by exporting JSON data to CSV (comma separated values) format and importing it into software NVivo 10 (QSR International) for further data storage, organization, and management. R was also used to visualize a word cloud associated with the highest frequency terms detected in Twitter content in order to better identify thematic categories in the data (“wordcloud” package).

We also manually coded the subset of tweets determined as relevant for NUPM for identification of any URLs/hyperlinks in tweets advertising online sale of prescription drugs. We used the NVivo NCapture feature to archive websites with hyperlinks in order to review content and determine if they were relevant to NUPM, acted as marketing affiliates, or directly sold prescription drugs online. A review of the risk status of illicit online pharmacy NUPM links was achieved by cross-referencing websites associated with hyperlinks using a database from the private Internet monitoring company, LegitScript, which provides information on websites likely in violation of applicable laws in one or more countries [25].

All tweets classified as NUPM-relevant that also had available latitude and longitude data were geocoded as individual points onto a contiguous map of the United States using ArcGIS version 10.1 (Esri). A publically available basemap was downloaded from the Esri website, and a scale of 1 centimeter:150 kilometers was applied. For the geographic coordinate system, we adopted the World Geodetic System 1984 standard. Zip code–level data on the number of individuals by age group was downloaded from the US Census Bureau, and the kernel density function was used to create a heatmap from this data, thereby displaying a gradient from blue to red for lower density of individuals between ages 15-19 to higher density of individuals in this age group, respectively. This was done to visualize the distribution of NUPM geocoded tweets to regions of the United States with a higher density of teens and young adults.

Results

Data Collection Results

We experienced no detectable interruptions in service during the 2-week study data period for our two separate Rstudio AWS instances. In total, our study data yielded 2,417,662 tweets, comprising 232,108 generic name tweets, 72.53% (n=168,355) of which were in English; and 2,185,554 street name tweets, 81.74% (n=1,786,626) of which were in English. Study data for the generic name instance yielded 1.44% (n=3351) that were geocoded for geographic location (a similar rate of location-enabled tweets compared to previous studies [26,27]), and 49.84% (n=115,685) that included URLs (comprising 22,174 unique URLs). Similarly, our street name NUPM instance yielded 1.39% (n=30,274) geocoded tweets and 43.51% (n=951,107) that included URLs (comprising 376,304 unique URLs). We describe preliminary data characteristics and their association with NUPM promotion, behavior, and access for both instances in the next section.

We note that during the study data period we received limit notices (limit count) for both instances collecting data from the public streaming API indicating that a filtered stream matched more tweets than the rate limit allowed to be delivered [28]. Twitter limit notices provide a total count of the number of undelivered tweets since the API connection was opened and can contextualize how representative filtered data collected from the Twitter public API is compared to the full Twitter firehose (which offers full and complete access to current and historic tweets). A total count of 247 tweets (corresponding to 0.1% of collected tweets) and 8327 (0.38%) in the generic and street name instances were reported, respectively.

Word Frequency and Association

Prior to conducting our content coding for study data, we first conducted text mining and analysis in R using the tm package. After generating a Term-Document Matrix, we analyzed Twitter NUPM content for word frequency and association. The most frequently observed words contained in the generic name dataset (frequency count includes all words used in a tweet including non-keyword terms) were for several drugs subject to abuse by youth and adolescents including: Valium (6.23% of all words observed, n=50,769, rank #1), Adderall (5.33%, n=43,426, rank #3), and Xanax (4.48%, n=36,486, rank #4). The relative high frequency of specific NUPM keywords indicates that even without using a separate process for iteratively filtering and coding of Twitter content, the original filters applied in this instance were more likely to yield a reasonable percentage of NUPM relevant tweets given the specificity of keywords.

Conversely, in the street name instance, we immediately noticed that non-specific street names for drugs (primarily “speed,” which refers to amphetamine; “reds,” which refers to barbiturates; and “benz,” which refers to benzedrine) were used in several thousands of tweets unrelated to drug abuse. The presence of tweets that were “noisy” (ie, prescription drug abuse slang terms that introduce extreme noise as they are used in more common words or return results about unrelated topics) in this instance necessitated further filtering of the dataset, pointing to the need for better construction of keyword filters during the initial data collection phase [29]. Based on these preliminary results, we conducted a separate filtering of the street name dataset to exclude the above three keywords that introduced noise and used this new dataset of 79,949 tweets during the content analysis process. The words with the highest frequency in the street name instance were “skippy” (slang for the drug Adderall/Ritalin, 2.73% of all words observed, n=12,453, rank #1), “yellows” (slang for barbiturates, 2.08%, n=9511, rank #2), “uppers” (slang for amphetamines, 2.08%, n=9498, rank #3), “barbs” (slang for barbiturates, 1.35%, n=6168, rank #4), and “oxy” (slang for OxyContin/Oxycodone, 1.15%, n=5269, rank #5).

Two separate word clouds were generated depicting the frequency of all words observed in the text of tweets analyzed for both instances (see Figures 2 and 3).

Figure 2. Word cloud for generic instance.

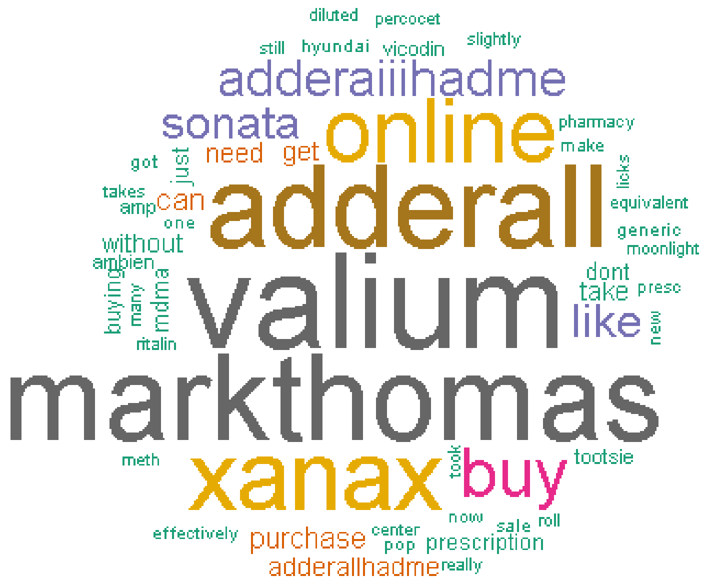


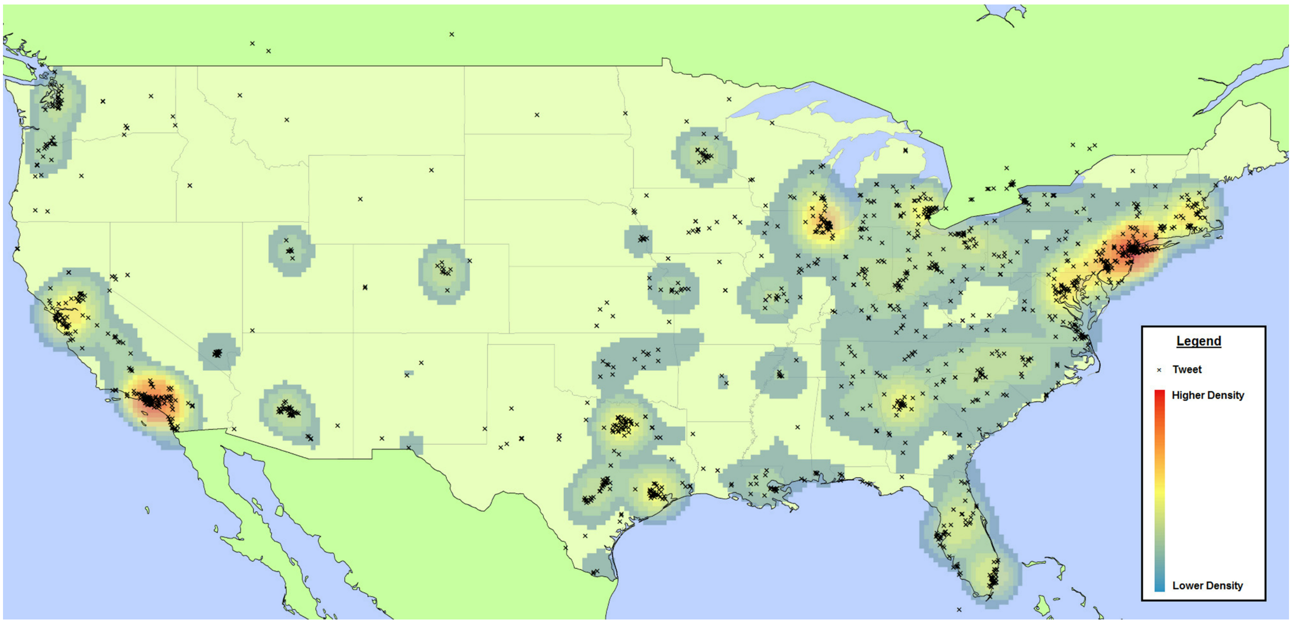
Figure 3. Word cloud for street instance.

Twitter NUPM Characteristics

In the next step, we applied our supervised machine learning content coding protocol in order to more specifically identify content relevant to NUPM. We used this process to filter out "false positive" tweets that were unrelated to NUPM but nevertheless were returned in filtered results as the keyword appeared somewhere in the text or metadata [22]. Coding of data between the 2 human coders achieved a high level of intercoder agreement for both subsamples of data from the two instances (all Cohen's kappas for characteristics reviewed were greater than .83 and had a mean score of *k*=.91). Supervised machine learning was conducted by defining a feature space of documents (tweets) with a unigram term-document matrix. The accuracy score of the classifier models of tweet relevance to NUPM, when evaluated by repeating 10-fold cross validation three times, was 94.5% for the generic name dataset and 93.5% for the street name dataset, while accuracy scores of favorability for the generic name data were 95.1% and 94.1% for the street name data. We also compared the performance of models built from different word-grouping units (unigram, bigram, trigram, and combination of 1-3 grams); however, larger grouping units did not add significant performance improvement compared to unigram analysis. We therefore used the model created with unigram text data for the final analysis.

Applying this classifier to the generic name dataset yielded a total of 135,776 tweets (81.72% of the English tweets with at least one keyword) that were predicted to be relevant to NUPM behavior or promotion. Conversely, false positives were detected in high frequency in the street name instance (eg, OXY, which is a stock listing symbol for the publicly traded company Occidental Petroleum Corporation and #uppers, which is a Twitter hashtag for users who engage in political discussion), with only 9817 (12.28%) tweets predicted as relevant to NUPM. Of the NUPM relevant tweets, an estimated 98.59% (n=133,863) and 78.76% (n=7732) favorably promoted NUPM behavior in the generic name and street name instances, respectively. Finally, after filtering both instances for NUPM-relevant tweets, 1.36% (n=1842) and 3.1% (n=308) were detected as geocoded, in the generic name and street name instances, respectively, with most of the tweets originating from users in California, Texas, and New York. In a map comparing individual geolocatable tweets with zip code–level kernel density of individuals aged 15-19, there was a positive relationship observed between NUPM tweets and areas with higher youth density (see Figure 4).

Figure 4. Geocoded NUPM relevant tweets by intensity and age demographics - United States.



Illegal Online Access to Prescription Drugs

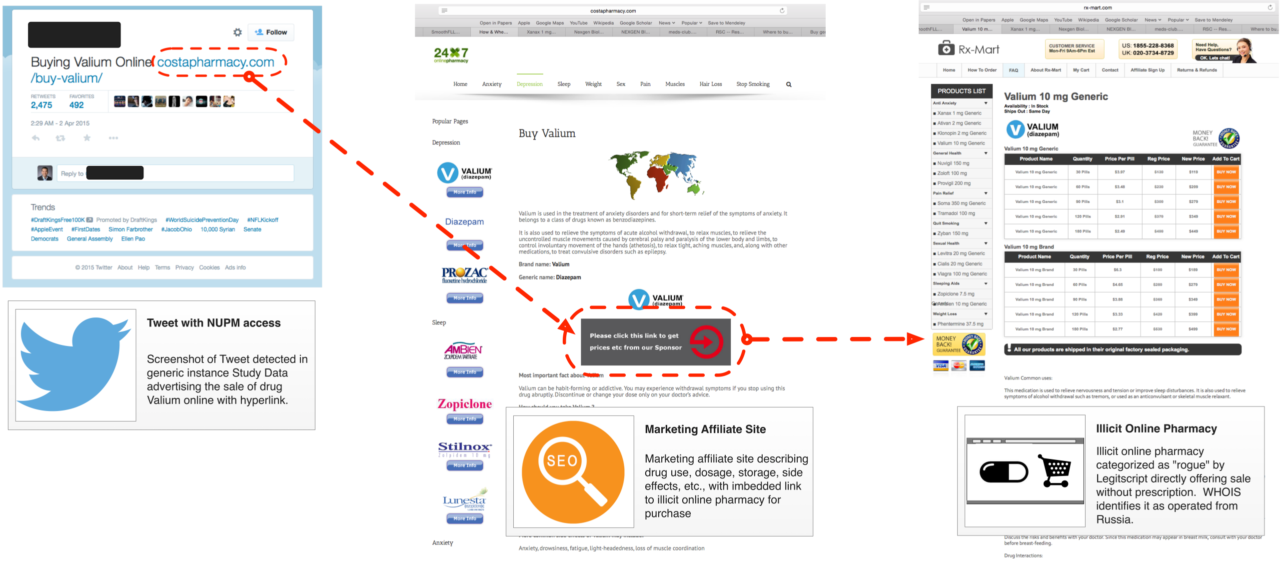
In the NUPM relevant tweets for the generic name dataset, there were 59,845 hyperlinks with 5039 unique URLs. The vast majority of these hyperlinks (88.54%, n=52,988) were tweeted at least 9 times. We then manually reviewed the websites connected to hyperlinks for content tweeted at least 9 times in order to determine if websites actually marketed or sold prescription drugs. The most frequently observed URL was for a purported online pharmacy using the name “CostaPharmacy.” It was mentioned in 75.72% (n=45,317) of all tweets with hyperlinks and on further inspection was identified as an online marketing affiliate site that promoted the online purchase of Valium (diazepam), a benzodiazepine type of medication used as a tranquilizer that is commonly abused. A total of 8171 different Twitter users tweeted or retweeted to this affiliate site that contained, within its website content, a direct link to an illegal online pharmacy. Importantly, the linked online pharmacy was categorized as “rogue” by LegitScript indicating that the website appears to be intentionally or knowingly violating applicable laws or regulations. On review of WHOIS records, it appears to be located in Russia (see Figure 5 for visualization of tweet, hyperlinks, and connection between websites). By counting the number of users following accounts that mentioned the URL, we estimate that content containing the hyperlink connected with this illegal online pharmacy was broadcast to over 250,000 total Twitter users within the 2-week study data period. We also observed a much smaller number of additional tweets with hyperlinks to NUPM access points that either advertised the sale of a prescription drug and then linked to another site or actually claimed to sell the drug directly to customers via their online storefront. All associated links to online pharmacy sites reviewed were categorized as “rogue.” Additionally, other links we observed promoted NUPM but were not associated with an online pharmacy. This included Twitter content with links to hip hop songs promoting the NUPM lifestyle, NUPM promotional items sold on eBay, and a tweet linking to a website selling nutritional supplements advertised as substitutes for NUPM drugs (see Table 2 for examples).

Table 2. Examples of different categories of NUPM tweets with website links.

Category	Tweet content	Description
Lifestyle: individual user	RT @[ANONYMIZED]: "I wanna try xanax" Me https://t.co/XpTErzo6W6	Retweet of individual user promoting drug abuse initiation with link to video where user shows pills in hand
Lifestyle: individual user (polydrug mention)	RT @[ANONYMIZED]: Prescription drugs, show me lovePercocets, Adderall, Xanny bars, get codeine involved	Tweet describing promotion of polydrug abuse with several different therapeutic classes
Lifestyle: music	RT @[ANONYMIZED]: Go do some xanax, crank this shithttps://t.co/9N72OXO0rs	Retweet with link to streaming hip hop song promoting prescription drug abuse behavior
Commercial: individual seller	Check out Xanax Pill Necklace http://t.co/IfeGqe01t4 via @eBay	Link to eBay seller account for a Xanax pill bar necklace that promotes NUPM lifestyle
Commercial: company	Over the counter? #Adderall?#Xanaxsubstitute #anxietynaturalremedy? #herebalremedy http://ow.ly/MsyqQ	Tweet linking to website selling nutritional supplements that are advertised as substitutes to Adderall and Xanax
Online pharmacy–related link	RT @[ANONYMIZED]: How to Buy Valium Online http://t.co/qkDY8ZJ08W	Retweet with link to marketing affiliate that included a link to an illicit online pharmacy

We note that in the street name dataset, no NUPM online access links were detected. This was determined by manually reviewing all 653 unique URLs included in the NUPM relevant tweets for the street name dataset, revealing that none of the links were associated with an online pharmacy.

Figure 5. Twitter NUPM content illicit online pharmacy relationship.



Tweet with NUPM access

Screenshot of Tweet detected in generic instance Study Data advertising the sale of drug Valium online with hyperlink.

Marketing Affiliate Site

Marketing affiliate site describing drug use, dosage, storage, side effects, etc., with imbedded link to illicit online pharmacy for purchase

Illicit Online Pharmacy

Illicit online pharmacy categorized as "rogue" by Legitscript directly offering sale without prescription. WHOIS identifies it as operated from Russia.

Overall Cost and Accessibility

Important in assessing the feasibility of this study is describing the overall cost of project implementation. Primarily, we attempted to conduct the data collection phase of this study with computational and software resources already available in the public domain. This included the use of the R programming language as a data collection and analysis platform, which utilized several software packages that were open access and free to use. Since the Twitter Streaming API collects data in real-time, contiguity of data primarily depends on the reliability of the network connectivity. We therefore used AWS services to ensure fidelity of our data collection (ie, service level guarantees of 99.95% uptime). Overall, AWS services to support this study were delivered at low cost, at a total expense of US $17 per instance with 40GB of storage space (excluding the AWS free Tier of 750 machine hours per month of t2.micro instances) for a total data collection period of approximately 3 months (analysis for the entire period of data collection is ongoing). Because fees for AWS services are variable based upon size of storage, CPU utilization, and memory size, by regularly transferring data to an external storage space one can also minimize the cost required for data storage directly on AWS.

Discussion

Principal Results

Based on our analysis of the study data collected, it appears that NUPM promotion via Twitter is occurring frequently for specific keywords associated with a drug’s chemical or generic name and yields large-scale datasets that require the appropriate combination of analytical and computing tools to appropriately assess potential behavioral risk factors and online access. In comparison, the more indistinct use of NUPM street or popular names that includes descriptive or non-specific terms introduced more noise and far fewer relevant results compared to the generic name dataset. From a data collection standpoint, our 2-week period of study data yielded over 2 million tweets from both instances, with a total of 6.02% (n=145,593) determined relevant to NUPM behavior or promotion based on our use of supervised machine learning for content analysis.

Our most concerning finding was that in the generic name instance, a significant percentage (33.37%, n=45,317) of the machine classified NUPM relevant content originated from a highly propagated live hyperlink to a marketing affiliate that provided direct access to prescription drugs of abuse through an illegal online pharmacy. The marketing affiliate site advertises the online sale of several drugs of abuse, including Ativan, Ambien, Lunesta, Valium, and Xanax, among other classes of prescription drugs. The tweets for this hyperlink varied slightly in content, but all blatantly advertised the sale of Valium (eg, “RT @[username anonymized]: Valium Online Without Prescription [URL]” or “RT @[username anonymized]: Where to Buy Valium Online [URL]”) and were retweeted by a large network of Twitter accounts potentially exposing hundreds of thousands of Twitter users to NUPM promotion and access. As a Drug Enforcement Agency Schedule IV controlled substance, the online sale of Valium and other drugs that carry the potential for abuse and dependence, is in direct violation of the Ryan Haight Online Pharmacy Consumer Protection Act, a federal law named after a San Diego Teen who lost his life after overdosing from illegally purchasing prescription pain killers online [1,6].

Finally, the relatively low count in undelivered tweets compared to the total tweets collected in each instance may indicate that our data collection methodology has a higher rate of completeness (99.9% for the generic name instance and 99.7% for the street name instance) and is a more representative sample compared to what is generally represented in the literature regarding the Twitter Public API [22]. It may also indicate that by creating separate virtual instances for data collection operating on different keyword filters and consumer keys/access tokens, our data collection process can avoid Streaming API rate limits (usually estimated at 1% sample of all tweets). Overall, the study supports findings from prior studies that have used the Twitter Streaming API to collect large amounts of Twitter content instead of licensing content from a third-party data reseller that may be cost-prohibitive to certain researchers (ie, the starting price for a GNIP Twitter dataset request is US $1250) [13,26,30,31].

Limitations

There are certain limitations to our study that impact the generalizability of results to NUPM behavior and promotion via Twitter. When hand coding tweets, human coders observed that some tweets were extremely short (ie, 1-2 words) and did not contain content that indicated that it promoted NUPM behavior or access, even though the keyword was contained in the tweet. These tweets were coded as non-relevant by human coders as they would likely be interpreted by Twitter users as non-relevant and were a very small percentage (5%) of overall human coded tweets. Additionally, the text of some tweets contained hyperlinks to images and other media associated with the tweet that helped contextualize the content/message or confirm promotion of NUPM behavior. Although human coders reviewed these images that were linked to these tweets, which aided in their interpretation of content classification, our machine learning algorithm was not able to analyze this media in subsequent machine classification prediction. We also estimated the number of Twitter users who potentially received the rogue online pharmacy Valium hyperlink by examining the followers_count field of the user statuses. Although this approach is limited in that the analysis can be performed only retrospectively and thus the followers_count may differ from the time the tweets were actually generated, the result is indicative of how a Twitter user network can play a role in promoting illicit online pharmacies to a broad base of users.

Finally, given the large scale of data collected per instance relative to the short time period examined, future content coding of NUPM Twitter data will likely need to be assisted with additional high-performance computational tools/services in order to make such a project scalable over a longer period of data collection. Further, more iterative rounds of human coding using Twitter data collected over a longer period of time with more diversity in Twitter users and content could help improve the machine learning process. Possible solutions to augment trained human coding include the use of crowdsourcing large networks of human coders (eg, Amazon Mechanical Turk workers) or the use of new content coding services in the cloud, such as those offered by the company DiscoverText, that offer cloud-based text analytics solutions targeted for analyzing social media data [32,33]. These tools and platforms have already been utilized in previous substance abuse studies and could be applied to future work analyzing a larger dataset of NUPM tweets [11,32-34]. We also did not filter for language as we are considering the possibility of using language information in future studies. Here, however, we content coded only English tweets and not non-English tweets (19.14%, n=462,681 of the dataset.) We also encountered a handful of “dead” links in the hyperlinks manually coded for association with an illicit online pharmacy, though we note in both cases this was an extremely small percentage of the total tweets collected over the study period.

Comparison With Prior Work

The few studies that have specifically examined the association between Twitter and NUPM have focused on testing the ability to illegally advertise illicit online pharmacy content via a fictitious Twitter account, qualitatively assessing tweets about prescription opioids, online social engagement between networks of prescription drug abusers that use Twitter, and Twitter use to promote drug abuse of Adderall among college students [7,10,19,35]. This study expands on previous studies to further explore how the Twitter environment can promote NUPM behavior and access by examining a broad set of prescription drug keywords associated with abuse by youth and adolescents. The study builds on previous research that has used different sets of prescription drug and common/slang keywords to filter and analyze Twitter data, as well as prior studies that have analyzed the content of hyperlinks detected in large-scale datasets of filtered tweets for other public-health related topics [16,36-38].

Conclusions

As youth and adolescents increasingly engage in online communities, social relationships, and conversations about NUPM via popular social media platforms such as Twitter, additional research is critical in order to leverage strategies of “infoveillance” to collect data needed to tailor future public health interventions attempting to combat prescription drug abuse among this vulnerable population [5,39]. Importantly, analysis of real-time data via Twitter, can help inform and contextualize traditional public health surveillance approaches collected through national surveys and also help proactively identify changing and emerging trends in prescription drug abuse behavior that are unique to the online environment. The study also identifies Twitter as a potential source for information illegally promoting the sale of controlled prescription drugs directly to consumers, which is a concerning observation given the inherent risk of abuse, dependency, and questionable authenticity of medicines provided by online pharmacies who are in violation of applicable law, including the US Ryan Haight Act. These results support renewed focus to better understand these understudied channels of NUPM promotion and needed commitment to develop technology-based tools, online health promotion activities, and public policy protecting youth and adolescents from prescription drug abuse online.

Acknowledgments

TK and TM received funding for this research from the Alliance for Safe Online Pharmacies (ASOP), a 501(c)(4) social welfare organization engaged in the issue of illicit online pharmacies, and greatly acknowledge this support. The funder had no role or input in the study.

Conflicts of Interest

TM is a non-compensated member of the ASOP academic advisory panel of ASOP, and both TK and TM received funding for this project from ASOP through a pilot research grant exploring prescription drug abuse risks online.

References

1. Mackey TK, Liang BA, Strathdee SA. Digital social media, youth, and nonmedical use of prescription drugs: the need for reform. J Med Internet Res 2013;15(7):e143 [FREE Full text] [doi: 10.2196/jmir.2464] [Medline: 23892156]
2. McCabe SE, West BT, Teter CJ, Boyd CJ. Co-ingestion of prescription opioids and other drugs among high school seniors: results from a national study. Drug Alcohol Depend 2012 Nov 1;126(1-2):65-70 [FREE Full text] [doi: 10.1016/j.drugalcdep.2012.04.017] [Medline: 22609061]
3. Inciardi JA, Surratt HL, Cicero TJ, Rosenblum A, Ahwah C, Bailey JE, et al. Prescription drugs purchased through the internet: who are the end users? Drug Alcohol Depend 2010 Jul 1;110(1-2):21-29 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.01.015] [Medline: 20227199]
4. Lenhard A. Pew Internet Research. 2015 Apr 09. Teens, Social Media & Technology Overview 2015 URL: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/ [accessed 2015-11-30] [WebCite Cache ID 6dQoCIXUh]
5. Stoové MA, Pedrana AE. Making the most of a brave new world: opportunities and considerations for using Twitter as a public health monitoring tool. Prev Med 2014 Jun;63:109-111. [doi: 10.1016/j.ypmed.2014.03.008] [Medline: 24632229]
6. Liang BA, Mackey T. Searching for safety: addressing search engine, website, and provider accountability for illicit online drug sales. Am J Law Med 2009;35(1):125-184. [Medline: 19534258]
7. Hanson CL, Cannon B, Burton S, Giraud-Carrier C. An exploration of social circles and prescription drug abuse through Twitter. J Med Internet Res 2013;15(9):e189 [FREE Full text] [doi: 10.2196/jmir.2741] [Medline: 24014109]
8. Duggan M, Ellison N, Lampe C, Lenhart A, Madden M. Pew Internet Research. 2015 Jan 09. Demographics of Key Social Networking Platforms Internet URL: http://www.pewinternet.org/2015/01/09/demographics-of-key-social-networking-platforms-2/ [accessed 2015-11-29] [WebCite Cache ID 6dPUw4mIf]
9. Ghosh DD, Guha R. What are we 'tweeting' about obesity? Mapping tweets with Topic Modeling and Geographic Information System. Cartogr Geogr Inf Sci 2013;40(2):90-102 [FREE Full text] [doi: 10.1080/15230406.2013.776210] [Medline: 25126022]
10. Hanson CL, Burton SH, Giraud-Carrier C, West JH, Barnes MD, Hansen B. Tweaking and tweeting: exploring Twitter for nonmedical use of a psychostimulant drug (Adderall) among college students. J Med Internet Res 2013;15(4):e62 [FREE Full text] [doi: 10.2196/jmir.2503] [Medline: 23594933]
11. Huang J, Kornfield R, Szczypka G, Emery SL. A cross-sectional examination of marketing of electronic cigarettes on Twitter. Tob Control 2014 Jul;23 Suppl 3:iii26-iii30 [FREE Full text] [doi: 10.1136/tobaccocontrol-2014-051551] [Medline: 24935894]
12. Bosley JC, Zhao NW, Hill S, Shofer FS, Asch DA, Becker LB, et al. Decoding twitter: Surveillance and trends for cardiac arrest and resuscitation communication. Resuscitation 2013 Feb;84(2):206-212 [FREE Full text] [doi: 10.1016/j.resuscitation.2012.10.017] [Medline: 23108239]
13. Myslín M, Zhu S, Chapman W, Conway M. Using twitter to examine smoking behavior and perceptions of emerging tobacco products. J Med Internet Res 2013;15(8):e174 [FREE Full text] [doi: 10.2196/jmir.2534] [Medline: 23989137]
14. O'Connor K, Pimpalkhute P, Nikfarjam A, Ginn R, Smith KL, Gonzalez G. Pharmacovigilance on twitter? Mining tweets for adverse drug reactions. AMIA Annu Symp Proc 2014;2014:924-933 [FREE Full text] [Medline: 25954400]
15. Freifeld C, Brownstein J, Menone C, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in twitter. Drug Saf 2014 May;37(5):343-350 [FREE Full text] [doi: 10.1007/s40264-014-0155-x] [Medline: 24777653]
16. Carbonell P, Mayer MA, Bravo A. Exploring brand-name drug mentions on Twitter for pharmacovigilance. Stud Health Technol Inform 2015;210:55-59. [Medline: 25991101]
17. Broniatowski DA, Paul MJ, Dredze M. National and local influenza surveillance through Twitter: an analysis of the 2012-2013 influenza epidemic. PLoS One 2013;8(12):e83672 [FREE Full text] [doi: 10.1371/journal.pone.0083672] [Medline: 24349542]
18. Dasgupta N, Freifeld C, Brownstein JS, Menone CM, Surratt HL, Poppish L, et al. Crowdsourcing black market prices for prescription opioids. J Med Internet Res 2013;15(8):e178 [FREE Full text] [doi: 10.2196/jmir.2810] [Medline: 23956042]
19. Mackey TK, Liang BA. Global reach of direct-to-consumer advertising using social media for illicit online drug sales. J Med Internet Res 2013;15(5):e105 [FREE Full text] [doi: 10.2196/jmir.2610] [Medline: 23718965]
20. NIDA. drugabuse. 2015 Oct. Commonly Abused Prescription Drugs Chart Internet URL: http://www.drugabuse.gov/drugs-abuse/commonly-abused-drugs-charts [accessed 2015-11-29] [WebCite Cache ID 6dPV9HuS2]
21. Yoon S, Elhadad N, Bakken S. A practical approach for content mining of Tweets. Am J Prev Med 2013 Jul;45(1):122-129 [FREE Full text] [doi: 10.1016/j.amepre.2013.02.025] [Medline: 23790998]
22. Kim AE, Hansen HM, Murphy J, Richards AK, Duke J, Allen JA. Methodological considerations in analyzing Twitter data. J Natl Cancer Inst Monogr 2013 Dec;2013(47):140-146. [doi: 10.1093/jncimonographs/lgt026] [Medline: 24395983]

23. Kuhn M. Building predictive models in R using the caret package. J Stat Soft 2008;28(5):-. [doi: 10.18637/jss.v028.i05]
24. Feinerer I, Hornik K, Meyer D. Text Mining Infrastructure in R. J Stat Soft 2008;25(5):1. [doi: 10.18637/jss.v025.i05]
25. LegitScript.com. Frequently Asked Questions URL: http://www.legitscript.com/about/faq/ [accessed 2015-11-30] [WebCite Cache ID 6dQolkzEc]
26. Nagar R, Yuan Q, Freifeld CC, Santillana M, Nojima A, Chunara R, et al. A case study of the New York City 2012-2013 influenza season with daily geocoded Twitter data from temporal and spatiotemporal perspectives. J Med Internet Res 2014;16(10):e236 [FREE Full text] [doi: 10.2196/jmir.3416] [Medline: 25331122]
27. Burton SH, Tanner KW, Giraud-Carrier CG, West JH, Barnes MD. "Right time, right place" health communication on Twitter: value and accuracy of location information. J Med Internet Res 2012;14(6):e156 [FREE Full text] [doi: 10.2196/jmir.2121] [Medline: 23154246]
28. Streaming message types. 2015. URL: https://dev.twitter.com/streaming/overview/messages-types [accessed 2015-11-30] [WebCite Cache ID 6dQorZ0Go]
29. Adrover C, Bodnar T, Huang Z, Telenti A, Salathé M. Identifying Adverse Effects of HIV Drug Treatment and Associated Sentiments Using Twitter. JMIR Public Health Surveill 2015 Jul 27;1(2):e7 [FREE Full text] [doi: 10.2196/publichealth.4488]
30. Zhang N, Campo S, Janz KF, Eckler P, Yang J, Snetselaar LG, et al. Electronic word of mouth on twitter about physical activity in the United States: exploratory infodemiology study. J Med Internet Res 2013;15(11):e261 [FREE Full text] [doi: 10.2196/jmir.2870] [Medline: 24257325]
31. McIver DJ, Hawkins JB, Chunara R, Chatterjee AK, Bhandari A, Fitzgerald TP, et al. Characterizing Sleep Issues Using Twitter. J Med Internet Res 2015;17(6):e140 [FREE Full text] [doi: 10.2196/jmir.4476] [Medline: 26054530]
32. Paul MJ, Dredze M. Discovering health topics in social media using topic models. PLoS One 2014;9(8):e103408 [FREE Full text] [doi: 10.1371/journal.pone.0103408] [Medline: 25084530]
33. Choundhury MD, Gamon M, Counts S, Horvitz E. Predicting Depression via Social Media. In: Seventh International AAAI Conference on Weblogs and Social Media. 2013 Presented at: ICWSM 2013; July 8-10, 2013; Boston, MA p. 128-137.
34. Choundhury MD, Counts S, Horvitz E. Social Media as a Measurement Tool of Depression in Populations. : ACM; 2013 Presented at: 5th Annual ACM Web Science Conference; May 2-4, 2013; Paris, France p. 47-56. [doi: 10.1145/2464464.2464480]
35. Shutler L, Nelson LS, Portelli I, Blachford C, Perrone J. Drug Use in the Twittersphere: A Qualitative Contextual Analysis of Tweets About Prescription Drugs. J Addict Dis 2015;34(4):303-310. [doi: 10.1080/10550887.2015.1074505] [Medline: 26364675]
36. Harris J, Moreland-Russell S, Choucair B, Mansour R, Staub M, Simmons K. Tweeting for and against public health policy: response to the Chicago Department of Public Health's electronic cigarette Twitter campaign. J Med Internet Res 2014;16(10):e238 [FREE Full text] [doi: 10.2196/jmir.3622] [Medline: 25320863]
37. Prochaska JJ, Pechmann C, Kim R, Leonhardt JM. Twitter=quitter? An analysis of Twitter quit smoking social networks. Tob Control 2012 Jul;21(4):447-449 [FREE Full text] [doi: 10.1136/tc.2010.042507] [Medline: 21730101]
38. Lee JL, DeCamp M, Dredze M, Chisolm MS, Berger ZD. What are health-related users tweeting? A qualitative content analysis of health-related users and their messages on twitter. J Med Internet Res 2014;16(10):e237 [FREE Full text] [doi: 10.2196/jmir.3765] [Medline: 25591063]
39. Nelson L, Meisel Z, Perrone J. Opportunities for Exploring and Reducing Prescription Drug Abuse Through Social Media. J Addict Dis 2015;34(2-3):178-184. [doi: 10.1080/10550887.2015.1059712] [Medline: 26075802]

Abbreviations

API: application programming interface
ASOP: Alliance for Safe Online Pharmacies
AWS: Amazon Web Services
JSON: JavaScript Object Notation
NUPM: non-medical use of prescription medication
SVM: Support Vector Machine (algorithm)
URL: uniform resource locator

Edited by G Eysenbach; submitted 18.09.15; peer-reviewed by C Giraud-Carrier; comments to author 09.10.15; revised version received 28.10.15; accepted 11.11.15; published 16.12.15.

Please cite as:
Katsuki T, Mackey TK, Cuomo R
Establishing a Link Between Prescription Drug Abuse and Illicit Online Pharmacies: Analysis of Twitter Data
J Med Internet Res 2015;17(12):e280
URL: http://www.jmir.org/2015/12/e280/
doi:10.2196/jmir.5144
PMID:26677966

Corrigenda and Addenda

Figure Correction: How Consumers and Physicians View New Medical Technology: Comparative Survey

Debra L Boeldt[1], PhD; Nathan E Wineinger[1], PhD; Jill Waalen[1], PhD; Shreya Gollamudi[1]; Adam Grossberg[2]; Steven R Steinhubl[1], MD; Anna McCollister-Slipp[1]; Marc A Rogers[2], PhD; Carey Silvers[2]; Eric J Topol[1], MD

[1]Scripps Translational Science Institute, Scripps Health, The Scripps Research Institute, La Jolla, CA, United States
[2]WebMD, New York, NY, United States

Corresponding Author:
Eric J Topol, MD
Scripps Translational Science Institute
Scripps Health
The Scripps Research Institute
3344 North Torrey Pines Ct
Suite 300
La Jolla, CA, 92037
United States
Phone: 1 858 554 5708
Fax: 1 858 546 9272
Email: etopol@scripps.edu

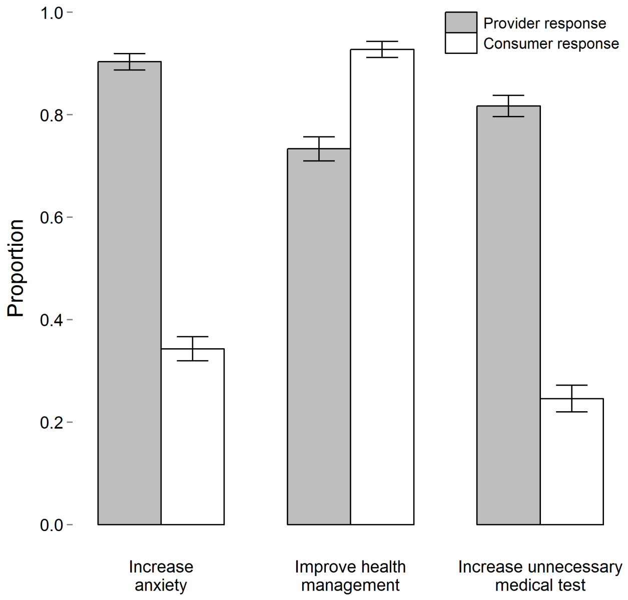
Related Article:

Correction of: http://www.jmir.org/2015/9/e215/

(J Med Internet Res 2015;17(12):e284) doi:10.2196/jmir.5150

In Figure 2 of the paper “How Consumers and Physicians View New Medical Technology: Comparative Survey” (J Med Internet Res 2015;17[9]:e215), authors erroneously inverted the bars indicating the proportion of people who believed access to electronic health records would increase anxiety in patients, improve health management, or increase the number of unnecessary medical tests. The originally published paper showed the proportion of respondents who answered “No” to the question. The corrected figure now displays the proportion of people who responded “Yes.” The online version of this JMIR paper has been updated with this figure, and a corrected version was sent to PubMed Central.

Figure 2. Proportion of responders who believed access to electronic health records would increase anxiety in patients, improve health management, or increase the number of unnecessary medical tests (error bars represent 95% confidence intervals).



Edited by G Eysenbach; submitted 21.09.15; this is a non–peer-reviewed article;accepted 22.09.15; published 18.12.15.

Please cite as:
Boeldt DL, Wineinger NE, Waalen J, Gollamudi S, Grossberg A, Steinhubl SR, McCollister-Slipp A, Rogers MA, Silvers C, Topol EJ
Figure Correction: How Consumers and Physicians View New Medical Technology: Comparative Survey
J Med Internet Res 2015;17(12):e284
URL: http://www.jmir.org/2015/12/e284/
doi:10.2196/jmir.5150
PMID:30578207

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