# interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 4 (2015), Issue 1 ISSN 1929-073X

## Contents

### Original Papers

### Review

## Short Paper

Original Paper

# Preferred Features of Oral Treatments and Predictors of Non-Adherence: Two Web-Based Choice Experiments in Multiple Sclerosis Patients

Paul Wicks[1], PhD; David Brandes[2], MD; Jinhee Park[3], PhD; Dimitri Liakhovitski[4], PhD; Tatiana Koudinova[4]; Rahul Sasane[3], PhD

[1]PatientsLikeMe, Cambridge, MA, United States
[2]Hope Neurology / Hope MS Center, Knoxville, TN, United States
[3]Novartis Pharmaceuticals Inc, East Hanover, NJ, United States
[4]GfK Custom Research, East Hanover, NJ, United States

**Corresponding Author:**
Paul Wicks, PhD
PatientsLikeMe
155 2nd Street
Cambridge, MA, 02141
United States
Phone: 1 617 499 4003
Fax: 1 866 850 6240
Email: pwicks@patientslikeme.com

## Abstract

**Background:** Oral disease modifying therapies (DMTs) for multiple sclerosis (MS) differ in efficacy, tolerability, and safety.

**Objective:** We sought to understand how these attributes impact patient preference and predicted DMT non-adherence among oral-naïve MS patients.

**Methods:** Adult MS patients from the “PatientsLikeMe” Web-based health data-sharing platform completed a discrete choice exercise where they were asked to express their preference for one of three hypothetical oral DMTs, each with a certain combination of levels of tested attributes. Another Web-based exercise tested a number of possible drivers of non-adherence, mainly side effects. Data from an MS clinic were used to adjust for sample bias. Respondents’ preferences were analyzed using Hierarchical Bayesian estimation.

**Results:** A total of 319 patients completed all questions. Most respondents were female (77.7%, 248/319) with mean age 48 years (SD 10). Liver toxicity was the attribute that emerged as the most important driver of patient preference (25.8%, relative importance out of 100%), followed by severe side effects (15.3%), delay to disability progression (10.7%), and common side effects (10.4%). The most important drivers of predicted non-adherence were frequency of daily dosing (17.4% out of 100%), hair thinning (14.8%), use during pregnancy (14.1%), severe side effects (13.8%), and diarrhea (13.0%).

**Conclusions:** Understanding the important concerns expressed by patients may help health care providers to understand and educate their patients more completely about these concerns. This knowledge may therefore improve both choices of appropriate therapy and adherence to therapy over time.

***(Interact J Med Res 2015;4(1):e6)*** doi:10.2196/ijmr.3776

**KEYWORDS**

multiple sclerosis; drug therapy; decision making; cross-sectional survey

## Introduction

Recent years have seen the introduction of a number of oral disease modifying therapies (DMTs) to the multiple sclerosis (MS) armamentarium, which supplements the earlier range of injectable and intravenous DMTs [1]. Adherence to injectable DMTs can be especially challenging for patients; although patients can often be resourceful in using coping strategies for injectable medications, reduced adherence remains an issue [2]. Three new oral DMTs have been approved for the treatment of

MS: fingolimod (Gilenya, Novartis, FDA Orange Book approval date September 21, 2010), teriflunomide (Aubagio, Genzyme/Sanofi, FDA Orange Book approval date September 12, 2012), and dimethyl fumarate (Tecfidera, Biogen Idec, FDA Orange book approval date March 27, 2013). These choices change the landscape considerably. Each oral DMT occupies a unique niche in terms of its quantified ability to delay MS rate of progression, reduce frequency of relapses, alter lesion burden visible on magnetic resonance imaging (MRI), and contribute to side effects or serious adverse events. Physicians and patients must become increasingly involved in tradeoffs between efficacy, convenience, and safety in decisions and, interestingly, each approaches these factors from different perspectives [3].

One technique that helps examine the tradeoffs made by patients is “consider jointly” analysis (“conjoint” analysis). In this approach, participants are asked to indicate their preference from one of several discrete choices for a number of consecutive hypothetical product profiles that differ by varying levels of selected attributes. Such techniques have been used in studies of other chronic illnesses with complex decisions and tradeoffs to be made about treatments, with several studies in serious illness [4]. Because conjoint analysis takes multiple attributes into account simultaneously, conjoint studies allow researchers to create more complex models of decision-making than responses on simple rating scales. For example, in a sample of men with prostate cancer, a conjoint analysis study found that men were willing to trade off life expectancy to be relieved of certain side effects of treatment, and that preferences differed by age [5]. Similar work has also been applied to tradeoffs between different treatment characteristics in acne vulgaris [6] and pain control in osteoarthritis [7].

In MS, Johnson et al [8] used a Web-based conjoint analysis to study risk tolerance in a group of over 600 MS patients including those who had previously used natalizumab (which has been associated with an increased risk of progressive multifocal leukoencephalopathy/PML [9]). In showing patients different levels of efficacy and risk for hypothetical treatments, they found the most important attributes influencing patient preference to be the effect of a treatment to slow disease progression (27% out of a possible 100%), closely followed by risk of PML (23%), liver failure (20%), leukemia (18%), and reduction in the frequency of relapse (12%) [9]. The authors concluded that patients were willing to make tradeoffs of risk in exchange for improved DMT efficacy.

Given the recent availability of oral DMTs and the complex factors underlying decisions about selecting one, we sought to explore the relative preferences of a sample of oral-naïve MS patients (who have never taken any oral DMTs) with regards to salient oral DMT attributes. Our primary objective was to quantify and rank these attributes. Given that injectable DMTs have a variety of barriers to adherence that are rendered irrelevant by oral DMTs, we also sought to quantify and rank attributes that might affect the likelihood of non-adherence to the oral DMTs.

## Methods

### Recruitment

Methods are reported in accordance with the Checklist for Reporting Results of Internet Surveys (CHERRIES) [10]. Over a 10-day period in July 2013, we fielded a cross-sectional survey to a population of existing oral-naïve members from the “PatientsLikeMe” Web-based health-data sharing platform who reported a diagnosis of multiple sclerosis, were aged 18 years or over, and living in the United States. Patients reporting prior use of oral DMTs were excluded to maintain oral naivety and infrequently prescribed DMTs (such as Extavia) were excluded from the study to avoid cells with small Ns for post-hoc analyses.

Members who sign up to PatientsLikeMe do so under the terms of use, which make clear they could be contacted for research; additional informed consent was collected for this voluntary study. Potential participants were selected on the basis of previously submitted profile data and were contacted via email. Participants were informed of the study sponsor, the objectives of the study, that it would take approximately 15 minutes to complete, and that a US $25 cash card incentive would be provided to those who completed the study. Consenting to the study took patients to a Web-based survey tool hosted by GfK Custom Research. To avoid missing or spurious data, all questions were mandatory to complete the survey, participants could not revise earlier answers, and unique URLs were used to avoid the risk of multiple completions or spurious data entries. Institutional review board (IRB) approval was granted from Western IRB.

### Survey Development

Survey measures included basic demographics, MS DMT history, patient-reported disease severity (MS Rating Scale revised, MSRS-R) [11], the MS Treatment Adherence Questionnaire (MSTAQ) [2], and the Beliefs about Medicines Questionnaire (BMQ) [12].

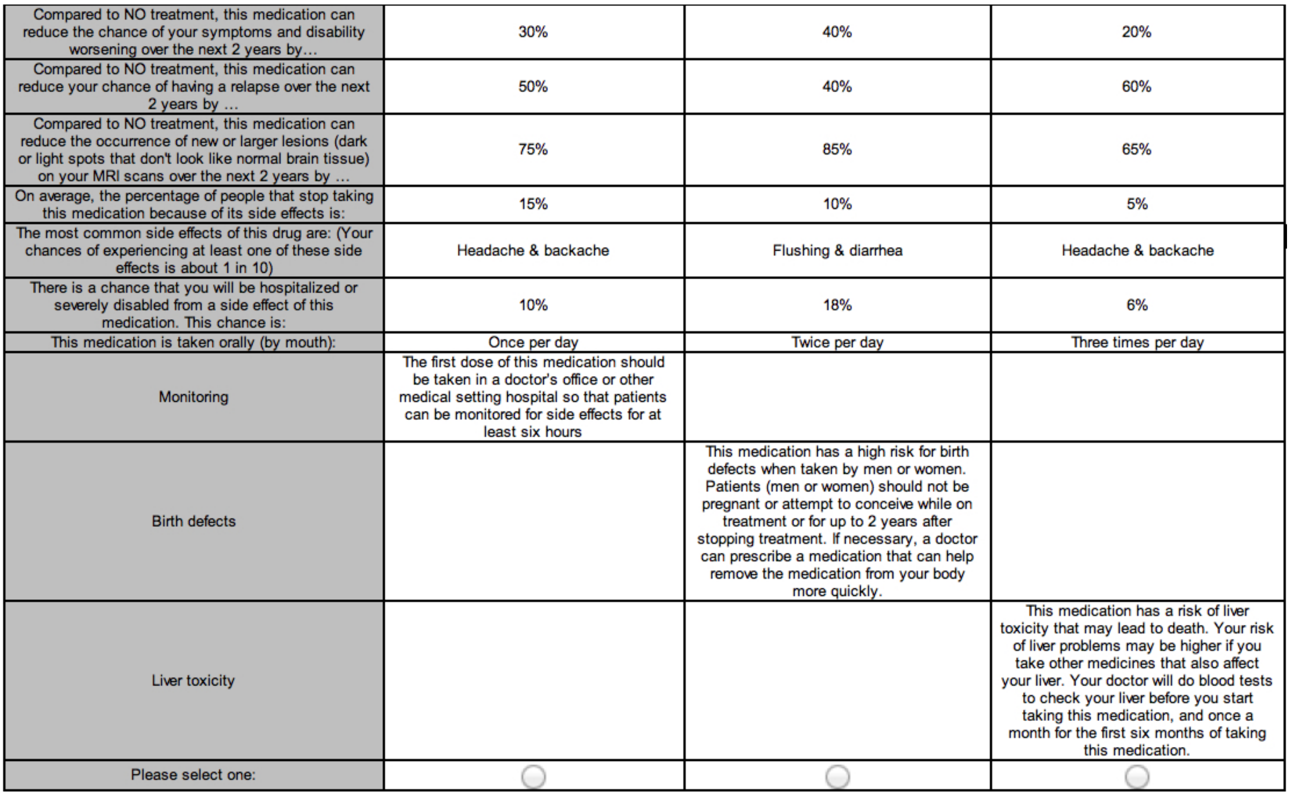
### Treatment Characteristics Preference Exercise (CA1)

We developed two conjoint analysis tasks in accordance with International Society For Pharmacoeconomics and Outcomes Research (ISPOR) guidance on use of conjoint analysis [11]. The first exercise (CA1) asked oral-naïve MS patients to repeatedly choose one of three hypothetical oral DMTs, each with a certain combination of levels of tested attributes (shown in Table 1). Example screenshots of the tasks are shown in Figures 1 and 2.

**Table 1.** Parameters and values for Conjoint Analysis Exercise 1 (CA1) ranked by order of importance to oral-naïve MS patients (n=319).

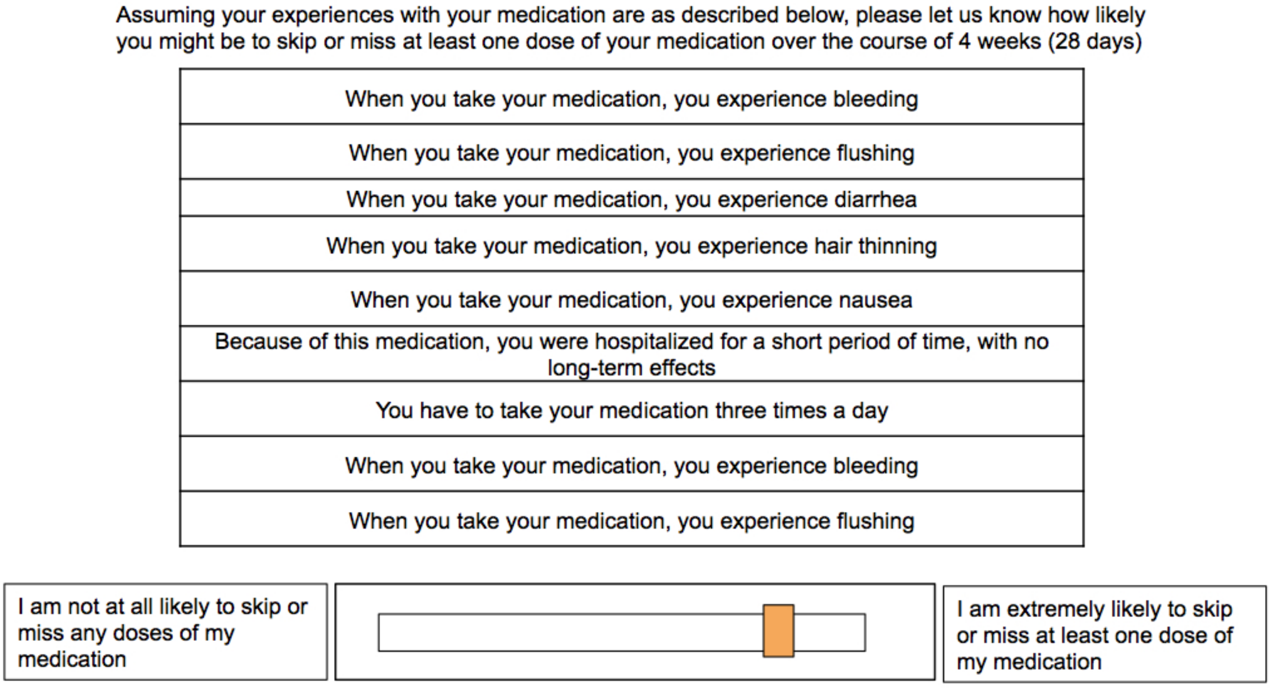
| Attribute | Description | Value Level 1 | Value Level 2 | Value Level 3 | Value Level 4 | Relative importance (out of 100%) |
|---|---|---|---|---|---|---|
| Liver toxicity | This medication has a risk of liver toxicity that may lead to death. Your risk of liver problems may be higher if you take other medicines that also affect your liver. Your doctor will do blood tests to check your liver before you start taking this medication, and once a month for the first six months of taking this medication. | Yes | No | – | – | 25.8% |
| Severe side effects | There is a _____ chance that you will be hospitalized or severely disabled from a side effect of this medication. | 6% | 10% | 14% | 18% | 15.3% |
| Delay the progression of disability | Compared to no treatment, this medication can reduce the chance of your symptoms and disability worsening over the next 2 years by… | 20% | 30% | 40% | – | 10.7% |
| Common side effects | The most common side effects of this drug are … Your chances of experiencing at least one of these side effects is about 1 in 10. | Headache, backache | Flushing, diarrhea | Hair thinning, nausea | – | 10.4% |
| Frequency of administration | This medication is taken orally (by mouth)… | Once per day | Twice per day | Three times per day | – | 9.5% |
| Reduce frequency of relapses | Compared to no treatment, this medication can reduce your chance of having a relapse over the next 2 years by… | 30% | 40% | 50% | 60% | 8.7% |
| Reduce changes on MRI | Compared to no treatment, this medication can reduce the occurrence of new or larger lesions (dark or light spots that don't look like normal brain tissue) on your MRI scans over the next 2 years by… | 65% | 75% | 85% | – | 6.7% |
| First dose monitoring | The first dose of this medication should be taken in a doctor's office or other medical setting hospital so that patients can be monitored for side effects for at least six hours. | Yes | No | – | – | 4.6% |
| Tolerability | On average, _____ of people stop taking this medication because of its side effects. | 5% | 10% | 15% | – | 4.5% |
| Birth defects | This medication has a high risk for birth defects when taken by men or women. Patients (men or women) should not be pregnant or attempt to conceive while on treatment or for up to 2 years after stopping treatment. If necessary, a doctor can prescribe a medication that can help remove the medication from your body more quickly. | Yes | No | – | – | 3.8% |

**Figure 1.** Example screenshot from Conjoint Analysis Exercise 1 (CA1) – Participants were asked “Of these three products, which would you be most likely to ask your physician to prescribe to you if these were the only options available?".



| | | | |
|---|---|---|---|
| Compared to NO treatment, this medication can reduce the chance of your symptoms and disability worsening over the next 2 years by… | 30% | 40% | 20% |
| Compared to NO treatment, this medication can reduce your chance of having a relapse over the next 2 years by … | 50% | 40% | 60% |
| Compared to NO treatment, this medication can reduce the occurrence of new or larger lesions (dark or light spots that don't look like normal brain tissue) on your MRI scans over the next 2 years by … | 75% | 85% | 65% |
| On average, the percentage of people that stop taking this medication because of its side effects is: | 15% | 10% | 5% |
| The most common side effects of this drug are: (Your chances of experiencing at least one of these side effects is about 1 in 10) | Headache & backache | Flushing & diarrhea | Headache & backache |
| There is a chance that you will be hospitalized or severely disabled from a side effect of this medication. This chance is: | 10% | 18% | 6% |
| This medication is taken orally (by mouth): | Once per day | Twice per day | Three times per day |
| Monitoring | The first dose of this medication should be taken in a doctor's office or other medical setting hospital so that patients can be monitored for side effects for at least six hours | | |
| Birth defects | | This medication has a high risk for birth defects when taken by men or women. Patients (men or women) should not be pregnant or attempt to conceive while on treatment or for up to 2 years after stopping treatment. If necessary, a doctor can prescribe a medication that can help remove the medication from your body more quickly. | |
| Liver toxicity | | | This medication has a risk of liver toxicity that may lead to death. Your risk of liver problems may be higher if you take other medicines that also affect your liver. Your doctor will do blood tests to check your liver before you start taking this medication, and once a month for the first six months of taking this medication. |
| Please select one: | ○ | ○ | ○ |

**Figure 2.** Example screenshot from Conjoint Analysis Exercise 2 (CA2) – Participants moved the yellow sliders between the two extreme values.



Assuming your experiences with your medication are as described below, please let us know how likely you might be to skip or miss at least one dose of your medication over the course of 4 weeks (28 days)

When you take your medication, you experience bleeding

When you take your medication, you experience flushing

When you take your medication, you experience diarrhea

When you take your medication, you experience hair thinning

When you take your medication, you experience nausea

Because of this medication, you were hospitalized for a short period of time, with no long-term effects

You have to take your medication three times a day

When you take your medication, you experience bleeding

When you take your medication, you experience flushing

I am not at all likely to skip or miss any doses of my medication

I am extremely likely to skip or miss at least one dose of my medication

## Non-Adherence Exercise (CA2)

The second exercise (CA2) showed patients just one hypothetical oral DMT but with varying levels of the values shown in Table 2 and a response task of a visual analogue scale they used to indicate how likely they might be to miss at least one dose of medication over the course of 4 weeks (28 days), with “I am not at all likely to skip or miss any doses of my medication” at one end and “I am extremely likely to skip or miss at least one dose of my medication” on the other. Both sets of attributes were identified through review of the clinical trial literature by two independent raters and consultation with a clinical expert (DB).

**Table 2.** Parameters and values for Conjoint Analysis Exercise 2 (CA2) ranked by order of importance to oral-naïve MS patients (n=319).

| Attribute | Value Level 1 | Value Level 2 | Value Level 3 | Relative importance (out of 100%) |
|---|---|---|---|---|
| Frequency of administration | Once per day | Twice per day | Three times per day | 17.4% |
| Side effects − hair thinning | Yes | No | − | 14.8% |
| Pregnancy − you or your partner become pregnant while taking your medication | Yes | No | − | 14.1% |
| Severe side effects − hospitalized for a short period of time, with no long-term effects | Yes | No | − | 13.8% |
| Side effects − diarrhea | Yes | No | − | 13.0% |
| Side effects − nausea | Yes | No | − | 10.7% |
| Side effects − backache | Yes | No | − | 8.8% |
| Side effects − headache | Yes | No | − | 4.5% |
| Side effects − flushing | Yes | No | − | 3.0% |

### Statistical Analysis

To help overcome the biases of a Web-based sample of patients from the PatientsLikeMe Web-based data platform, we used sample weights to adjust the proportions of the sample for groups using benchmarks from the Partners Northeast MS Center in the United States [13]. A SAS iterative proportional fitting macro created the weights by adjusting all target variables simultaneously. Extreme weights were then trimmed to reduce the influence of extreme outliers in the weighted results and improve weighting efficiency. Trimmed weights are shown in Table 3 (weighting was associated with a design effect of 1.27, or 79% weighting efficiency).

Data were analyzed using hierarchical Bayesian estimation. We estimated hierarchical Bayesian models in Sawtooth Software (Orme, USA) that uses a specific Monte Carlo Markov Chain Algorithm called the “Metropolis Hastings Algorithm”. At the end of the estimation, each level of each attribute is assigned a numeric value (“part worth” or “utility”) that reflects how much this level is valued by the respondent. Relative importance of the attributes derives from conjoint analysis and is based on the utilities. The attribute importance values add up to 100% for each conjoint analysis exercise. For each attribute, a difference between the highest and the lowest utility is calculated, and the relative importance is obtained by dividing that difference by the sum of the differences for all attributes. The target sample size (N=300) was determined using a power analysis assuming three comparable products for each decision task, a 6% margin of error, and a desired 95% confidence interval. Most published conjoint analysis studies have a sample size between 100 and 300 respondents, and our proposed sample size is consistent with guidance in the methodology literature [4].

## Results

From 1790 invited MS patients, 327 completed all questions. During data cleaning, responses from 8 patients were not analyzed (2 duplicates, 2 violated inclusion criteria, 4 for “straight-lining” answers), leaving a total of 319 (17.8% response rate, 319/1790). Most respondents were female (77.7%, 248/319) with mean age 48 years (SD 10) and a 10-year self-reported history of MS with 70.8% (226/319) reporting a diagnosis of relapsing remitting MS. Table 3 shows the raw unweighted sample demographics as well as the transformed sample following weighting to more closely resemble a representative MS population. All conjoint analysis data referred to from this point comes from the weighted sample used for analysis.

The most frequently impaired aspect of function on the MSRS-R was walking, with 53.6% (171/319) of patients moderately-severely impacted on this item, followed by 36.4% (116/319) experiencing sensory issues, 26.0% (83/319) cognitive issues, and 27.0% (87/319) bowel or bladder disturbance.

Most patients (61.0%, 195/319) were taking a DMT at the time of survey, with the most frequent being glatiramer acetate (29.5%, 94/319, Copaxone, subcutaneous daily injection, TEVA) followed by natalizumab (18.8%, 60/319, Tysabri, monthly intravenous transfusion, Biogen Idec), interferon beta-1a (9.1%, 29/319, Avonex, weekly intramuscular injection, Biogen Idec), and interferon beta-1b (8.8%, 28/319, Rebif, 3x weekly subcutaneous injection, EMD Serono). Approximately two-thirds (61.1%, 195/319) were taking a DMT at the time of survey, slightly higher than the Sonya Slifka longitudinal study, which found that 50.0% of patients were “currently” (in 2000-2001) using a DMT but similar to the reported 62.2% that had taken a DMT at some point in their disease [14].

The BMQ showed that overall, most patients (67.1%, 214/319) worried about long-term effects of their DMTs, with a higher proportion endorsing this sentiment among the patients who had missed at least one dose in the past 28 days (77.6% agreed or strongly agreed, 52/67, compared to those that hadn’t (57.8%, 74/128, $\chi^2_{195}$= 9.780, $P$=.04).

Among the 195 patients using a DMT, most patients self-injected their DMT (69.2%, 135/195), with 22.1% (43/195) reporting that someone else helped them with the injection most or all of the time, and 8.7% (17/195) only half of injections or just a few times. Among the 156 patients who reported using an injectable DMT, 59.6% (93/156) used an auto-injector exclusively, 32.1% (50/156) injected manually, and 6.7% (13/195) used a mixture of both. Overall, ease of use with current treatment was relatively high, with 89.7% (175/195) of DMT users saying their treatment was either "easy to use" or only "a little hard" to use. Similarly, satisfaction was quite high, with most users reporting "moderately", "very", or "completely" satisfied (88.7%, 173/195) and only 11.3% reporting they were either "a little satisfied" or "not satisfied at all" (11.3%, 22/195).

About a third of DMT users (34.3%, 67/195) reported missing at least one dose in the previous 28 days, with the most common reasons being "did not feel like taking my medication" (35.8%, 24/67 reported as a "moderately" or "extremely" important factor), "memory problems" (26.9%, 18/67), and "tired of taking my medication" (28.4%, 19/67). Patients who reported missing a dose of their medication in the past 28 days were not significantly different than those who did not on their MSRS-R outcome ($t_{193}$=1.730, $P$=.09).

**Table 3.** Sample demographics before and after weighting (n=319).[a]

| | | Unweighted sample frequencies (%) | Benchmark (Partners MS Center) | Weighted sample frequencies (%) |
|---|---|---|---|---|
| **Gender** | | | | |
| | Female | 248 (77.7%) | 74.8% | 238 (74.6%) |
| | Male | 71 (22.3%) | 25.2% | 82 (25.7%) |
| **Age, years** | | | | |
| | 18-38 | 36 (11.2%) | 25.0% | 76 (23.8%) |
| | 39-46 | 59 (18.5%) | 25.0% | 79 (24.7%) |
| | 47-62 | 186 (58.3%) | 40.0% | 132 (41.3%) |
| | 63+ | 38 (11.9%) | 10.0% | 33 (9.7%) |
| **Race** | | | | |
| | White | 283 (88.7%) | 92.4% | 292 (91.5%) |
| | Black | 18 (5.6%) | 4.6% | 17 (5.3%) |
| | Other | 18 (5.6%) | 3.0% | 10 (3.1%) |
| **MS subtype** | | | | |
| | Relapsing-Remitting | 240 (75.2%) | 70.2% | 225 (70.5%) |
| | Primary Progressive | 25 (7.8%) | 6.2% | 18 (5.6%) |
| | Secondary Progressive | 47 (14.7%) | 22.6% | 72 (22.6%) |
| | Progressive Relapsing | 7 (2.1%) | 1.0% | 4 (1.2%) |
| **Highest education** | | | | |
| | High school graduate or less | 45 (14.1%) | 19.4% | 62 (19.4%) |
| | Some college | 115 (36.1%) | 54.8% | 171 (53.6%) |
| | College graduate or more | 159 (49.8%) | 25.8% | 86 (26.9%) |

[a] Percentages may not add up to 100% due to rounding.

## Treatment Characteristics Preference Exercise (CA1)

Results from the CA1 preference exercise showed that potential for liver toxicity was the most important factor (Table 1, 25.8% as a measure of relative importance out of 100%) in hypothetical DMT selection followed by severe side effects (15.3%), delay to progression of disability (10.7%), common side effects (10.4%), and mode of administration (9.5%). Reducing the frequency of relapses (8.7%) and reducing changes on MRI (6.7%) were less important in driving preferences, as were requirement for a first-dose monitoring period (4.6%), tolerability (4.5%), and risk of birth defects (3.8%).

## Non-Adherence Exercise (CA2)

The CA2 non-adherence exercise (Table 2) found that the most important determinant of self-reported non-adherence to a hypothetical DMT was frequency of daily administration (17.4%), hair thinning (14.8%), becoming pregnant (14.1%), severe side effects (13.8%), and diarrhea (13.0%). Nausea (10.7%), backache (8.8%), headache (4.5%), and flushing (3.0%) emerged as less important drivers of non-adherence. A tradeoff simulator was built for CA2, which allows prediction of non-adherence for any combination of the relevant values of these parameters. For instance, in the worst-case scenario of a thrice-daily dosing of an oral DMT that causes hair thinning, has severe side effects, diarrhea, nausea, backache, headache, and flushing, the model predicted a 78% likelihood of missing at least one dose over the course of 4 weeks for an average patient. In the best-case scenario, a DMT with none of the tested side effects and once-daily dosing, the model predicted 15% likelihood of at least one missed dose. In simulations, the frequency of daily dosing had the largest incremental impact on adherence.

## Discussion

### Principal Findings

In this Web-based survey of oral DMT naïve patients, we found that liver toxicity, severe side effects, and common side effects were the most salient attributes driving DMT preference, with efficacy, frequency of dosing, and first dose monitoring less important. However, we also found that the frequency of dosing and specific side effects, such as hair thinning, might have an influence on patients' predicted non-adherence to taking an oral DMT.

The landscape of MS is currently undergoing a transformation from self-injectable and intravenous DMTs to a wider range of delivery routes including oral agents [15]. Our findings suggest the primacy of serious adverse events like liver toxicity may be major drivers of patient preference. All three of the recently approved oral DMTs have some kind of hepatotoxic profile identified in their phase III trials, such as elevated alanine aminotransferase tests three times the normal range in 6-7% of patients taking dimethyl fumarate (with no reported hepatic failure) [16,17], 7-19% for fingolimod [18,19], and 7-12% for teriflunomide [20,21]. However, in this last case due to the known properties of leflunomide (which metabolizes to teriflunomide) in rheumatoid arthritis, the drug was issued with an FDA "black box" warning [22] for severe liver injury including fatal liver failure, requiring liver monitoring at least monthly for 6 months after treatment initiation. Patient concern over a similar risk appeared to be the major driver of patient preference in our hypothetical conjoint analysis and patients actually prescribed the drug might benefit from extra assurance that monitoring should identify any issues that arise.

Our results contrast with those of Johnson et al, who found a stronger preference for slowing disability (27%) than avoiding side effects, although the serious adverse event PML (23%) was a frequently endorsed driver of preference, as was the potential for liver failure (20%) [8]. The different findings between our two studies may reflect the different sample chosen, as at least 42% of Johnson et al's sample had already taken natalizumab (Tysabri) at some point, which may reflect a higher risk tolerance. By contrast, only 21% of our sample were taking Tysabri, and about a third were not taking any form of DMT. Although speculative, it is also possible that reports of PML in the MS community have drawn attention to the fact that even rare adverse events can occur with serious consequences [23].

In the current study, patients expressed greater preference for product profiles with fewer serious side effects and fewer common side effects relative to those with higher levels of efficacy. The clinical trial literature suggests that serious adverse events were reported among 17-18% of patients taking dimethyl fumarate in trials [16,17], followed by 14-16% over 12 weeks [20] taking teriflunomide (29-36% over long term use [21]), with the fewest among patients taking fingolimod (7-10%) [18,19]. However, it is worth nothing that oral medications do not carry the injection-related profile of side effects such as injection site pain or erythema [17].

In agreement with Johnson et al, we found a higher preference for delaying the progression of disability over reducing the frequency of relapses. Establishing the former in typical clinical trials is much harder than the latter, requiring longer and larger studies. It may be that patients better understand the concept of progressive disability than they do of "relapses", which are highly unpredictable and may be complex to disentangle from disability progression or pseudo-exacerbations.

As a complex condition involving many tradeoffs, there is increasing interest in the use of conjoint analysis techniques in various aspects of decision-making in MS and supporting MS patients to make better-informed decisions based on their personal treatment preferences [24]. Shingler et al describe the use of conjoint techniques to identify patient preferences for characteristics of self-injection devices [25] and found that a treatment's efficacy mattered more to patients than ease of use to administer it, with technological features like medication reminders having relatively low importance. Over half of our sample were using an auto-injector for their DMT and reporting a good level of satisfaction, begging the question of whether patients with a relatively high level of satisfaction will appreciate as much of a difference between a self-injected DMT and an oral as earlier cohorts of patients who did not have the benefit of auto-injectors.

Medication non-adherence is a known issue in MS and a variety of solutions has been proposed to study this important issue. The largest (N=2648) and most rigorous study in this disease, the Global Adherence Project (GAP), found that 25% of patients were non-adherent to therapy, with memory being a major issue [26]. Although studies of injectable DMTs have found a number of issues related to route of administration or site injection reactions, it is unlikely that memory issues or treatment fatigue will be addressed solely by a move to oral DMT therapy. Although greater convenience would be anticipated, it is worth noting that many MS patients are already reporting a high level of satisfaction particularly due to the use of auto-injectors. One possible downside to oral DMTs is the absence of adherence-tracking technology that can be built into auto-injector devices, although systems such as the Proteus Raisin System might address such challenges in future [27].

### Limitations

This study had several limitations, including the examination of only hypothetical product profiles in a cross-sectional manner and what patients said they would choose in an artificial setting rather than the behavior they would actually exhibit. However, conjoint analysis may be considered ecologically valid because individuals are used to making decisions from among multiple varying choices on a daily basis [28]. In terms of study design, all conjoint studies suffer from a conceptual bias in that the questions they seek to address naturally constrain patient choice in a way that may not reflect the real world. For example, in the current study we asked patients to select between one of three oral MS DMTs, when they might have preferred self-injectable DMTs or second-line DMTs such as natalizumab—patients face a wide array of potential DMT options [1]. Another conjoint analysis limitation is accurately conveying clinical endpoints and the concept of risk such as percentage changes; we attempted to use endpoints that were commonly used in the MS community (and vetted these with a clinician) but it is certainly feasible that cognitive biases or comprehension issues limited full understanding.

As an online community, PatientsLikeMe users exhibit biases relative to other clinical samples such as MS patients at a specialist MS center [13] including being younger or more likely to be female; we sought to address this through sample weighting. While weighting reduces the bias in results, it does increase the variance of the results, resulting in a decreased statistical sensitivity to detect differences between groups. Our sample contained a relatively high proportion of Copaxone users, which might affect the results by including a larger set of patients using a lower-risk drug. However, inspection of the conjoint analysis results suggest that there were no major differences between the preferences of current Copaxone users relative to those who have never used it, with differences in CA1 importance levels of just 1% or less. The sample also included a relatively high proportion of patients not taking any DMT. It is unclear why these participants were not taking a DMT but we felt it was important to include the results as they represent a proportion of the population that might one day stand to benefit from a DMT if their expectations can be met. Given the heterogeneity of experiences with medication adherence, defining “non-adherence” as a single missed dose in the past 28 days is overly simplistic, but as the number in this group was relatively low, any further subdivision would have lead to very low N’s for statistical analysis.

Given the self-reported nature of the site, we have no evidence to confirm that members saying they have been diagnosed with MS have actually done so, nor that this diagnosis was accurate. Research is underway to more systematically address this limitation, but we intentionally targeted for recruitment those members who had been active on the site in the preceding 120 days rather than recruiting a new sample through advertising, so the likelihood that a given member would have signed up more than three months previously and maintained an active but fake account just on the chance of later gaining a survey incentive seems low. We also cleaned the dataset for evidence of straight-lining or duplicate entry. It is conceivable that a subset of users might have gone on to have their diagnosis changed to another condition; however, these limitations are shared in common with many other studies that use mailing lists or Web-based recruitment techniques and held as a common caveat. In summary, we believe that all study methodologies have their own set of limitations but that Web-based techniques have the advantage of adapting to address these through iterative software upgrades.

Future areas of research might include studying how patients starting an oral DMT report making that decision and what their medication adherence is like long-term. There is an inherent assumption that moving from injectable to oral DMTs should produce improved adherence but this is yet to be tested rigorously.

In attempting to select the best of these therapeutic options for each patient, a balance must be struck of efficacy, safety, tolerability, adherence, potential need for monitoring, and cost effectiveness [29]. It is possible that the use of decision aids that personally tailor an individual patient’s attitudes to risk and lifestyle preferences, supported by quantitative data abstracted from the clinical literature, could prove a useful tool.

## Conclusions

Oral-naïve MS patients identified liver toxicity and serious side effects as the most significant determinants of DMT selection while high frequency of daily dosing and certain side effects appear to be the most important barriers to DMT adherence. The use of conjoint analysis could be helpful in the development of new decision aids to help patients and clinicians navigate their many choices of DMT.

## Acknowledgments

We are grateful to the members of PatientsLikeMe for volunteering to share their experiences, and to Kate Sulham and Darcy Krzynowek for contributions to the project. No writing assistance was used in this study.

This study was funded by Novartis Pharmaceuticals Inc.

## Conflicts of Interest

PW is an employee of PatientsLikeMe, owns stock/options in the company, and is a named inventor on patents filed by PatientsLikeMe. The PatientsLikeMe R&D team has received research support from Abbvie, Acorda, Actelion, Amgen, Astra-Zeneca, Avanir, Biogen, Genzyme, Johnson & Johnson, Merck, Novartis, Sanofi and UCB.
DB reports the following disclosures: Speakers Panel/Advisory Board: Acorda, Avanir, BiogenIdec, Genzyme, Novartis, Questcor, Serono, Teva; Research Support: Biogen Idec, Teva.
DL and TK are employees of GfK Custom Research.
JP and RS are employees of Novartis, and hold stock/options in the company.

## References

1. Castro-Borrero W, Graves D, Frohman TC, Flores AB, Hardeman P, Logan D, et al. Current and emerging therapies in multiple sclerosis: a systematic review. Ther Adv Neurol Disord 2012 Jul;5(4):205-220 [FREE Full text] [doi: 10.1177/1756285612450936] [Medline: 22783370]

2. Wicks P, Massagli M, Kulkarni A, Dastani H. Use of an online community to develop patient-reported outcome instruments: the Multiple Sclerosis Treatment Adherence Questionnaire (MS-TAQ). J Med Internet Res 2011;13(1):e12 [FREE Full text] [doi: 10.2196/jmir.1687] [Medline: 21266318]
3. Heesen C, Kleiter I, Nguyen F, Schäffler N, Kasper J, Köpke S, et al. Risk perception in natalizumab-treated multiple sclerosis patients and their neurologists. Mult Scler 2010 Dec;16(12):1507-1512. [doi: 10.1177/1352458510379819] [Medline: 20826527]
4. Marshall D, Bridges JF, Hauber B, Cameron R, Donnalley L, Fyie K, et al. Conjoint analysis applications in health - How are studies being designed and reported?: An update on current practice in the published literature between 2005 and 2008. Patient 2010 Dec 1;3(4):249-256. [doi: 10.2165/11539650-000000000-00000] [Medline: 22273432]
5. Sculpher M, Bryan S, Fry P, de Winter P, Payne H, Emberton M. Patients' preferences for the management of non-metastatic prostate cancer: discrete choice experiment. BMJ 2004 Feb 14;328(7436):382 [FREE Full text] [doi: 10.1136/bmj.37972.497234.44] [Medline: 14751919]
6. Kellett N, West F, Finlay AY. Conjoint analysis: a novel, rigorous tool for determining patient preferences for topical antibiotic treatment for acne. A randomised controlled trial. Br J Dermatol 2006 Mar;154(3):524-532. [doi: 10.1111/j.1365-2133.2005.07047.x] [Medline: 16445786]
7. Hauber AB, Arden NK, Mohamed AF, Johnson FR, Peloso PM, Watson DJ, et al. A discrete-choice experiment of United Kingdom patients' willingness to risk adverse events for improved function and pain control in osteoarthritis. Osteoarthritis Cartilage 2013 Feb;21(2):289-297 [FREE Full text] [doi: 10.1016/j.joca.2012.11.007] [Medline: 23182815]
8. Johnson FR, Van Houtven G, Ozdemir S, Hass S, White J, Francis G, et al. Multiple sclerosis patients' benefit-risk preferences: serious adverse event risks versus treatment efficacy. J Neurol 2009 Apr;256(4):554-562. [doi: 10.1007/s00415-009-0084-2] [Medline: 19444531]
9. Clifford DB, De Luca A, DeLuca A, Simpson DM, Arendt G, Giovannoni G, et al. Natalizumab-associated progressive multifocal leukoencephalopathy in patients with multiple sclerosis: lessons from 28 cases. Lancet Neurol 2010 Apr;9(4):438-446. [doi: 10.1016/S1474-4422(10)70028-4] [Medline: 20298967]
10. 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]
11. Bridges J, Hauber A, Marshall D, Lloyd A, Prosser L, Regier D, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 2011 Jun;14(4):403-413. [doi: 10.1016/j.jval.2010.11.013] [Medline: 21669364]
12. Horne R, Weinman J, Hankins M. The beliefs about medicines questionnaire: The development and evaluation of a new method for assessing the cognitive representation of medication. Psychology & Health 1999 Jan;14(1):1-24. [doi: 10.1080/08870449908407311]
13. Bove R, Secor E, Healy BC, Musallam A, Vaughan T, Glanz BI, et al. Evaluation of an online platform for multiple sclerosis research: patient description, validation of severity scale, and exploration of BMI effects on disease course. PLoS One 2013;8(3):e59707 [FREE Full text] [doi: 10.1371/journal.pone.0059707] [Medline: 23527256]
14. Minden S, Hoaglin D, Jureidini S, Hadden L, Frankel D, Komatsuzaki Y, et al. Disease-modifying agents in the Sonya Slifka Longitudinal Multiple Sclerosis Study. Mult Scler 2008 Jun;14(5):640-655. [doi: 10.1177/1352458507086463] [Medline: 18566028]
15. Perumal J, Khan O. Emerging disease-modifying therapies in multiple sclerosis. Curr Treat Options Neurol 2012 Jun;14(3):256-263. [doi: 10.1007/s11940-012-0173-x] [Medline: 22426573]
16. Gold R, Kappos L, Arnold DL, Bar-Or A, Giovannoni G, Selmaj K, DEFINE Study Investigators. Placebo-controlled phase 3 study of oral BG-12 for relapsing multiple sclerosis. N Engl J Med 2012 Sep 20;367(12):1098-1107. [doi: 10.1056/NEJMoa1114287] [Medline: 22992073]
17. Fox RJ, Miller DH, Phillips JT, Hutchinson M, Havrdova E, Kita M, CONFIRM Study Investigators. Placebo-controlled phase 3 study of oral BG-12 or glatiramer in multiple sclerosis. N Engl J Med 2012 Sep 20;367(12):1087-1097. [doi: 10.1056/NEJMoa1206328] [Medline: 22992072]
18. Cohen JA, Barkhof F, Comi G, Hartung HP, Khatri BO, Montalban X, TRANSFORMS Study Group. Oral fingolimod or intramuscular interferon for relapsing multiple sclerosis. N Engl J Med 2010 Feb 4;362(5):402-415. [doi: 10.1056/NEJMoa0907839] [Medline: 20089954]
19. Kappos L, Radue EW, O'Connor P, Polman C, Hohlfeld R, Calabresi P, FREEDOMS Study Group. A placebo-controlled trial of oral fingolimod in relapsing multiple sclerosis. N Engl J Med 2010 Feb 4;362(5):387-401. [doi: 10.1056/NEJMoa0909494] [Medline: 20089952]
20. O'Connor P, Wolinsky JS, Confavreux C, Comi G, Kappos L, Olsson TP, TEMSO Trial Group. Randomized trial of oral teriflunomide for relapsing multiple sclerosis. N Engl J Med 2011 Oct 6;365(14):1293-1303. [doi: 10.1056/NEJMoa1014656] [Medline: 21991951]
21. Confavreux C, Li DK, Freedman MS, Truffinet P, Benzerdjeb H, Wang D, Teriflunomide Multiple Sclerosis Trial Group. Long-term follow-up of a phase 2 study of oral teriflunomide in relapsing multiple sclerosis: safety and efficacy results up to 8.5 years. Mult Scler 2012 Sep;18(9):1278-1289 [FREE Full text] [doi: 10.1177/1352458512436594] [Medline: 22307384]

22. Rukovets O. Neurology Today. 2012. FDA Approves Oral MS Drug Teriflunomide URL: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm319277.htm [WebCite Cache ID 6WQplnuf1]
23. Fox RJ, Rudick RA. Risk stratification and patient counseling for natalizumab in multiple sclerosis. Neurology 2012 Feb 7;78(6):436-437. [doi: 10.1212/WNL.0b013e318245d2d0] [Medline: 22282644]
24. Wilson L, Loucks A, Bui C, Gipson G, Zhong L, Tang J, et al. Patient preferences for attributes of disease modifying therapies: Results of a choice based conjoint analysis. Value in Health 2013 May;16(3):A107. [doi: 10.1016/j.jval.2013.03.505]
25. Shingler SL, Swinburn P, Ali S, Perard R, Lloyd AJ. A discrete choice experiment to determine patient preferences for injection devices in multiple sclerosis. J Med Econ 2013 Aug;16(8):1036-1042. [doi: 10.3111/13696998.2013.811079] [Medline: 23730944]
26. Devonshire V, Lapierre Y, Macdonell R, Ramo-Tello C, Patti F, Fontoura P, et al. The Global Adherence Project (GAP): a multicenter observational study on adherence to disease-modifying therapies in patients with relapsing-remitting multiple sclerosis. Eur J Neurol 2011 Jan;18(1):69-77. [doi: 10.1111/j.1468-1331.2010.03110.x] [Medline: 20561039]
27. May M. Aiding adherence: five approaches to following prescriptions. Nat Med 2010 May;16(5):504. [doi: 10.1038/nm0510-504] [Medline: 20448562]
28. Ryan M. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. Soc Sci Med 1999 Feb;48(4):535-546. [Medline: 10075178]
29. Gasperini C, Ruggieri S, Mancinelli CR, Pozzilli C. Advances in the treatment of relapsing-remitting multiple sclerosis - critical appraisal of fingolimod. Ther Clin Risk Manag 2013;9:73-85 [FREE Full text] [doi: 10.2147/TCRM.S17426] [Medline: 23483794]

## Abbreviations

**BMQ:** Beliefs about Medicines Questionnaire
**CA:** conjoint analysis
**DMT:** disease-modifying therapy
**FDA:** Food and Drug Administration
**MRI:** magnetic resonance imaging
**MS:** multiple sclerosis
**MSRS:** MS Rating Scale revised
**MSTAQ:** MS Treatment Adherence Questionnaire
**PML:** progressive multifocal leukoencephalopathy

*Edited by G Eysenbach; submitted 12.08.14; peer-reviewed by G Mancardi; comments to author 05.12.14; accepted 21.01.15; published 05.03.15.*

*Please cite as:*
*Wicks P, Brandes D, Park J, Liakhovitski D, Koudinova T, Sasane R*
*Preferred Features of Oral Treatments and Predictors of Non-Adherence: Two Web-Based Choice Experiments in Multiple Sclerosis Patients*
*Interact J Med Res 2015;4(1):e6*
*URL: http://www.i-jmr.org/2015/1/e6/*
*doi:10.2196/ijmr.3776*
*PMID:25749630*

Original Paper

# Patterns of Technology Use in Patients Attending a Cardiopulmonary Outpatient Clinic: A Self-Report Survey

Rebecca T Disler[1*], BN, BSc, PGDipAdvPractice, MSc(Research); Sally C Inglis[1*], BN, BHSc(Hons), PhD, FAHA; Phillip J Newton[1*], BN(Hons), PhD; David C Currow[2*], BMed, PhD; Peter S Macdonald[3*], BMBS, PhD; Allan R Glanville[4*], MBBS, MD; DorAnne Donesky[5*], ANP-BC, PhD; Virginia Carrieri-Kohlman[5*], FAAN, RN, PhD; Patricia M Davidson[6*], FAAN, RN, MEd, PhD

[1]Centre for Cardiovascular and Chronic Care, Faculty of Health, University of Technology, Sydney, Broadway, Australia
[2]Discipline Palliative and Supportive Services, Flinders University, Adelaide, Australia
[3]St Vincent's Hospital and Victor Chang Cardiac Research Institute, Sydney, Australia
[4]St Vincent's Hospital, Sydney, Australia
[5]Physiological Nursing, University of California, San Francisco, CA, United States
[6]Department of Acute and Chronic Care, Johns Hopkins University, Baltimore, MD, United States
[*]all authors contributed equally

**Corresponding Author:**
Rebecca T Disler, BN, BSc, PGDipAdvPractice, MSc(Research)
Centre for Cardiovascular and Chronic Care
Faculty of Health
University of Technology, Sydney
PO Box 123
Broadway, NSW 2040
Australia
Phone: 61 427466732
Fax: 61 427466732
Email: rebecca.disler@uts.edu.au

## Abstract

**Background:** Self-management education for cardiopulmonary diseases is primarily provided through time-limited, face-to-face programs, with access limited to a small percentage of patients. Telecommunication tools will increasingly be an important component of future health care delivery.

**Objective:** The purpose of this study was to describe the patterns of technology use in patients attending a cardiopulmonary clinic in an academic medical center.

**Methods:** A prevalence survey was developed to collect data on participant demographics (age in years, sex, and socioeconomic status); access to computers, Internet, and mobile phones; and use of current online health support sites or programs. Surveys were offered by reception staff to all patients attending the outpatient clinic.

**Results:** A total of 123 surveys were collected between March and April 2014. Technological devices were a pervasive part of everyday life with respondents engaged in regular computer (102/123, 82.9%), mobile telephone (115/117, 98.3%), and Internet (104/121, 86.0%) use. Emailing (101/121, 83.4%), researching and reading news articles (93/121, 76.9%), social media (71/121, 58.7%), and day-to-day activities (65/121, 53.7%) were the most common telecommunication activities. The majority of respondents reported that access to health support programs and assistance through the Internet (82/111, 73.9%) would be of use, with benefits reported as better understanding of health information (16/111, 22.5%), avoidance of difficult travel requirements and time-consuming face-to-face appointments (13/111, 18.3%), convenient and easily accessible help and information (12/111, 16.9%), and access to peer support and sharing (9/111, 12.7%). The majority of patients did not have concerns over participating in the online environment (87/111, 78.4%); the few concerns noted related to privacy and security (10/15), information accuracy (2/15), and computer literacy and access (2/15).

**Conclusions:** Chronic disease burden and long-term self-management tasks provide a compelling argument for accessible and convenient avenues to obtaining ongoing treatment and peer support. Online access to health support programs and assistance was reported as useful and perceived as providing convenient, timely, and easily accessible health support and information.

Distance from the health care facility and a lack of information provision through traditional health sources were both barriers and enablers to telehealth. This is particularly important in the context of a cardiopulmonary clinic that attracts patients from a large geographical area, and in patients who are most likely to have high health care utilization needs in the future. Telecommunication interfaces will be an increasingly important adjunct to traditional forms of health care delivery.

***(Interact J Med Res 2015;4(1):e5)*** doi:10.2196/ijmr.3955

**KEYWORDS**

chronic obstructive pulmonary disease; chronic disease; self-management; self-care; telemedicine, eHealth; mHealth

## Introduction

The increasing burden of non-communicable diseases, such as heart and respiratory disease, is placing increasing pressure on global health systems [1,2]. The incidence and cumulative burden of these chronic progressive disorders is accentuated through population aging [1]. The prevalence of chronic heart failure (CHF) is 23 million worldwide with an overall prevalence of 2-3% of the population in the United States and Europe [2-5]. The global prevalence of chronic obstructive pulmonary disease (COPD) is estimated at 65 million and COPD is now responsible for 5% of all deaths globally [6,7]. Despite optimal pharmacological and medical treatments, individuals with COPD and CHF continue to experience high symptom burden, most commonly dyspnea and fatigue [8-12]. Both COPD and CHF are frequent causes of hospitalization and require self-management strategies [8-12]. The economic costs of COPD are approximately US$ 40 billion annually and this financial burden will only increase [8-12]. Daily symptoms, poor physical functioning, progressive social isolation, and caregiver burden contribute to this disease burden [6,7,13].

The burden of non-communicable diseases extends over time and the life course [14,15]. Self-management education for people with COPD and CHF, including symptom management strategies, exercise, and reinforcement of activity and medication adherence, are primarily provided through pulmonary and cardiac rehabilitation and heart failure specific disease management programs [16-20]. These interventions are commonly episodic, of short duration, and available only to a small percentage of individuals [16,17], with access limited by functional debilitation associated with chronic illness [21,22]. Although discrete disease management strategies are an integral element of evidence-based care, it is increasingly apparent that there are some symptom management issues that are germane across chronic conditions [23]. Self-management support should be targeted through multiple modes of delivery with a broad-based symptom focus [24]. Although chronic conditions such as COPD and CHF have received greater attention from the medical community over the past decades, the burden of disease at an individual level is less well recognized [25].

The most effective and economically sustainable approaches to support patients with chronic illnesses such as COPD and CHF, beyond acute exacerbations, require future investigation [13,18,24,26-28]. The disease burden and long-term tasks of self-management that confront patients are a compelling argument for accessible and convenient avenues to obtaining ongoing treatment and peer support [29-31]. Access to Web-based health information and support is well established in the United States with a recent report noting that over 50% of adults aged over 65 years use the Internet [28,32-35]; however, internationally, use is not so widespread [28,30-36]. In the Australian context, studies have explored Web-based health interventions, but there is limited information as to the patterns of technology use in this particular patient group [37].

Although technology access is challenging for some older adults who are the most burdened with chronic conditions, communication tools have become a critical component of health care delivery [29-31]. Rapid advances in tools that provide instant access to health information and rich resources for self-care have already created paradigm shifts in health consumer attitudes about their health and health care [28-31]. The evolution of eHealth (health care delivery through Internet and telehealth communications for surveillance, health promotion, and symptom or disease management) and the introduction of mHealth (monitoring, personal digital assistants, and other wireless devices) are markedly altering the collaboration and interaction between consumers, health providers, and institutions [38-40]. Asynchronous forms of health interaction, such as through email or discussion boards, allow individuals to receive self-management and condition support by posing questions to their provider without having to establish a formal face-to-face consultation (synchronous interaction) [36,41]. These converging factors will shape the development and testing of future interventions aimed at improving health outcomes and reducing costs across chronic illnesses. The new generation of empowered health consumers will expect that health care systems accommodate their changing needs and preferences for how they receive care, including access to evidence-based therapies [28-31].

In order to determine the future feasibility of Internet-based health care delivery, the reported prevalence study was undertaken to describe the patterns of technology use in patients attending a cardiopulmonary clinic in an academic medical center. The cardiopulmonary patients responding to this survey have provided a sample of those individuals most likely to have high health care utilization needs in the future; it is important to take this initial step in understanding whether these consumers are technology ready [42-44].

## Methods

### Objective

The objective of the study was to describe the prevalence and patterns of technology use in patients attending a cardiopulmonary outpatient clinic through a self-report survey.

### Recruitment

Patients attending a cardiopulmonary outpatient clinic at an academic medical center were invited to participate in this anonymous survey. All patients attending the cardiopulmonary outpatient clinic were eligible to participate.

The cardiopulmonary clinic is located within an academic medical center and provides services for patients with a variety of conditions including COPD, CHF, advanced lung disease, heart transplantation, and pulmonary hypertension. This clinic is a central referral setting for surrounding regional areas; consequently, individuals travel from all areas within the state to access specialist treatment.

### Instrument

A prevalence survey was developed in consultation with experts in the field of chronic illness and Internet-based health care delivery. The survey was presented in four sections with 11 questions used to capture information on participant demographics (age in years, sex, socioeconomic status); access and use of computers, Internet, and mobile telephones; and currently accessed health support sites. Socioeconomic status was described using the Australian socioeconomic indexes for areas (SEIFA) [45]. These indices summarize “the relative socioeconomic advantage and disadvantage of areas using data from the Census of Population and Housing” and are reported through Australian area postcode (area zip code). Indices are based on a number of variables including employment, private and rented occupied housing, family makeup, and highest qualification, to name a few [45]. Nominal tick boxes and free text short answer questions were used to collect responses. Respondents were able to give multiple answers to appropriate nominal and free text questions, noted by “please tick all that apply”.

The survey was piloted for 1 week in the cardiopulmonary clinic and 10 surveys were checked for completion and coherence with the questions asked prior to continuing with data collection. There was limited missing data in this initial phase, however, “Please turn over” was added to the bottom of the page for ease. No other adjustments were required. The final survey contained four sections with 11 questions and took approximately 5-10 minutes to complete (see Multimedia Appendix 1).

### Data Collection

Surveys were offered by reception staff to all patients attending the outpatient clinic at appointment registration. Surveys on clipboards were also placed on tables within the waiting area for patients to complete as they wished. Participation was voluntary with hard copy surveys completed and placed anonymously in a sealed submission box within the waiting area.

### Data Analysis

Descriptive statistics were used to analyze all aspects of the survey data.

### Ethical Issues

Ethical approval was provided by the collaborating academic institution and clinical site; approval numbers LHR/13/SVH/5 and 2012-149A. Participation was voluntary and anonymous.

## Results

### Respondents

A total of 123 surveys were collected between March and April 2014. Approximately 543 patients attended appointments at the cardiopulmonary clinic during the study period, resulting in an overall response rate of 22.7%. The overwhelming majority of respondents completed the survey questions in full. This took into consideration respondents who answered “no” to regular computer or Internet who were precluded from completing particular subsequent questions; all previous responses from these respondents were included in the descriptive statistics. All 123 respondents answered questions in regards to gender, with more females (72/123, 58.5%) noted to have completed the surveys than males (51/123, 41.5%). Age was reported in 118 of 123 (95.9%) surveys with median respondent age of 56 years (range 18-77), and 52.5% (62/118) of respondents aged between 50 and 64 years. All respondents noted their area zip code and from this just under half (55/123, 44.7%) of the respondents were considered to live in middle socioeconomic areas with under one-third coming from low socioeconomic areas (32/123, 26.0%) and under one-third living in high socioeconomic areas (36/123, 29.3%) (Table 1).

**Table 1.** Respondent demographic characteristics.

| Descriptive characteristics | | n (%) |
|---|---|---|
| **Gender (n=123)** | | |
| | Male | 51 (41.5) |
| | Female | 72 (58.5) |
| **Age, years (n=118)** | | |
| | Median (range) | 56 (18-77) |
| | Under 50 | 42 (35.6) |
| | 50-64 | 62 (52.5) |
| | Over 65 | 14 (11.9) |
| **Socioeconomic indexes for areas based on postcode (SEIFA), Australia, 2011 [a] (n=123)** | | |
| | Low income (Deciles 1 and 2) | 32 (26.0) |
| | Middle income (Deciles 3 to 8) | 55 (44.7) |
| | High income (Deciles 9 and 10) | 36 (29.3) |

[a] Australian Bureau of Statistics. Socioeconomic indexes for areas: robustness, diversity within larger areas, and the new geography standard Commonwealth of Australia 2012; ABS Catalogue no. 1351.0.55.038.

## Computer Use

All 123 respondents answered questions related to computer use with the majority of respondents engaged in regular computer use (102/123, 82.9%), defined as more than four times per week. The overwhelming majority had access to a device at home (118/123, 95.9%) mainly in the form of a laptop (91/123, 77.1%); however, over half additionally had access to a desktop (60/123, 50.8%) and a tablet (60/123, 50.8%). Fewer than half the respondents had access to a computer at work for personal use (58/123, 47.2%) and in most cases this access was a desktop computer (44/58, 75.9%). There was no marked difference in computer use across age groups or gender; however, respondents who came from lower socioeconomic areas (32/123, 26.0%) noted less regular computer use (24/32, 75%) compared with other groups (47/55, 85% in middle and 31/36, 86% in high socioeconomic groups) (Table 2).

**Table 2.** Questions relating to access and use of technology.

| Access to technology | n (%)[a] |
|---|---|
| **Regular computer use (n=123)** | |
| Yes | 102 (82.9) |
| No | 21 (17.1) |
| **Regular computer use, “yes”, by age group, years (n=118)** | |
| Under 50 (n=42) | 33 (78.6) |
| 50 – 65 (n=62) | 53 (85.5) |
| 65 and over (n=14) | 11 (78.6) |
| **Regular computer use, “yes”, by socioeconomic area [b] (n=123)** | |
| Low socioeconomic area (n=32) | 24 (75.0) |
| Middle socioeconomic area (n=55) | 47 (85.5) |
| High socioeconomic area (n=36) | 36 (86.1) |
| **Access to a computer device at home (n=118)** | |
| Desktop | 60 (50.8) |
| Laptop | 91 (77.1) |
| Tablet | 60 (50.8) |
| **Access to a computer device through work (n=58)** | |
| Desktop | 44 (75.9) |
| Laptop | 31 (53.4) |
| Tablet | 17 (29.3) |
| **Regular Internet use (n=121)** | |
| Yes | 104 (86.0) |
| No | 17 (14.0) |
| **Regular Internet use, “yes”, by age group, years (n=116)** | |
| Under 50 (n=41) | 37 (90.2) |
| 50 – 65 (n=62) | 52 (83.9) |
| 65 and over (n=13) | 11 (84.6) |
| **Regular Internet use, “yes”, by socioeconomic area [b] (n=121)** | |
| Low socioeconomic area (n=32) | 25 (78.1) |
| Middle socioeconomic area (n=54) | 48 (88.9) |
| High socioeconomic area (n=35) | 31 (88.6) |
| **Mode of Internet access at home (n=120)** | |
| Yes | 113 (94.2) |
| No | 7 (5.8) |
| **If yes to home Internet access, (n=110)** | |
| Wireless | 61 (55.5) |
| Broadband | 34 (30.9) |
| Cable/DSL/fiber | 9 (8.2) |
| Dial-up | 2 (1.8) |
| Unsure | 4 (3.6) |
| **Mode of Internet access outside the home (n=120)** | |
| Yes | 81 (67.5) |

| Access to technology | | n (%)[a] |
|---|---|---|
| | No | 39 (32.5) |
| **If yes to Internet access outside the home (n=93)** | | |
| | At work | 62 (66.7) |
| | Via public wireless | 58 (62.4) |
| | Via smartphone | 31 (33.3) |
| | Via friend’s place | 24 (25.8) |
| | Via Internet café | 10 (10.8) |
| **Key Internet activities (n=121)** | | |
| | Emailing | 101 (83.4) |
| | Browsing, researching, reading news articles | 93 (76.9) |
| | Social media | 71 (58.7) |
| | Day to day activities (shopping, banking, and browsing) | 65 (53.7) |
| | Browsing for health information | 56 (46.3) |
| | Skype or video calls | 36 (29.8) |
| **Access to a mobile phone (n=117)** | | |
| | Yes | 115 (98.3) |
| | No | 2 (1.7) |
| **Key mobile phone activities (n=115)** | | |
| | Phone calls | 111 (96.5) |
| | Sending texts | 100 (86.9) |
| | Internet browsing | 62 (53.9) |
| | Checking and sending emails | 57 (49.5) |
| | Other (playing games, social media, apps) | 6 (5.3) |

[a] Multiple responses to questions were accepted in free text questions and respondents were instructed to “tick all that apply” when responding to nominal questions. In this context, the sum of percentages will be more than 100%.

[b] Australian Bureau of Statistics. Socioeconomic indexes for areas: robustness, diversity within larger areas, and the new geography standard Commonwealth of Australia 2012; ABS Catalogue no. 1351.0.55.038

## Mobile Phone Use and Activities

The majority (117/123, 95.1%) of respondents answered questions related to mobile phone use and activities, with all but two respondents reporting that they used a mobile phone (115/117, 98.3%). Phone calls (111/115, 96.5%) and sending texts (100/115, 86.9%) were the two main activities carried out using a mobile phone. Over half of the respondents additionally used their phone for Internet browsing (62/117, 53.0%) and half for checking and sending emails (57/115, 49.5%) (Table 2).

## Internet Use and Activities

The majority (121/123, 98.4%) of respondents answered questions related to Internet use and activities, with the majority reporting regular Internet use (104/121, 86.0%). Internet use did not differ across age or gender; however, similar to computer use, those from lower socioeconomic areas had a reduced regular Internet use (25/32, 78%). Internet in the home setting was accessed by 94.2% (113/120) of respondents and in the main this was through wireless (61/110, 55.5%) or through broadband access (34/110, 30.9%). The majority of respondents also reported access to the Internet outside the home (81/120, 67.5%) and this was accessed either at work in line with computer access above (62/93, 67%) or through public wireless (58/93, 62%). A further third of individuals additionally had access to the Internet through smartphones (31/93, 33%) and others had access through a friend’s home (24/93, 26%) and Internet cafes (10/93, 11%) (Table 2).

The main activities undertaken through an Internet platform were reported in 121 of 123 (98.4%) of respondents with emailing (101/121, 83.4%), browsing, researching, and reading news articles (93/121, 76.9%), accessing social media (71/121, 58.7%), and day-to-day activities including online shopping, banking, and general browsing (65/121, 53.7%), as the most common. Just under half of respondents (56/121, 46.3%) used the Internet to browse health information and under a third (36/121, 29.8%) for Skype and video calling (Table 2). More female respondents noted that they used the Internet for both social media (female 50/69, 72% vs male 21/47, 45%,) and daily activities including online shopping banking and browsing (female 42/69, 61%, vs male 23/47, 49%) compared with their male counterparts. In regard to socioeconomic status, those from higher income areas showed a higher rate of email (32/34, 94% vs 43/51, 84% and 26/31, 84% in middle and lower socioeconomic groups respectively), research and reading the news (30/34, 88% vs 39/51, 76% in middle and 39/51, 77% in lower socioeconomic groups), and accessing health information through the Internet (21/34, 62% vs 23/51, 45% in middle and 12/31, 39% in lower socioeconomic groups). Those respondents from middle socioeconomic areas were more likely to access social media (37/51, 73%) compared with the other groups (17/34, 50% in higher and 17/31, 55% in lower socioeconomic groups). Respondents from lower socioeconomic areas were additionally less likely to Skype (6/31, 19% vs 12/34, 35% in higher and 18/51, 35% in lower socioeconomic groups) or engage in daily online activities, such as shopping, banking, and browsing (11/31, 35% vs 23/34, 86% in higher and 31/51, 61% in lower socioeconomic groups) (Table 2).

### The Potential for Web-Based Support and Information

The majority of respondents (111/123, 90.2%) answered questions in relation to access, concerns, and use of technology-based health websites. The majority answered that they would find it useful to have access to support programs and assistance with health problems through the Internet (82/111, 73.9%). Respondents between the ages of 50 to 65 years had a slightly higher positive response to this (45/54, 83%) compared with those in the under 50 years group (25/40, 63%) and the over 65 years group (8/12, 67%). Interestingly, those from higher socioeconomic areas were less likely to respond positively to finding benefit from online support and information, with only 67% (22/33) responding “yes”, compared with 75% (21/28) in lower socioeconomic areas and 78% (39/50) in middle socioeconomic areas. The majority of respondents gave reasons as to why they would access online support (71/111, 63.9%) with the main reasons being: better able to understand health information and condition management (16/71, 23%), avoid difficult travel requirements and time-consuming face-to-face appointments (13/71, 18%), and have convenient and easily accessible help and information (12/71, 17%). Nine (13%) of 71 reported “the more help the better” or words to that effect, and nine (13%) of 71 noted the benefit of peer support and sharing. It is also important to note that six (8%) of 17 respondents wrote that online information would address the difficulty they experienced in accessing information from their health providers and a further six (8%) of 71 noted that they would be able to get up-to-date advice on management and treatments (Table 3).

**Table 3.** Questions regarding online access, concerns, and currently used sites.

| Online access, concerns and currently accessed sites | | n (%)[a] |
|---|---|---|
| **Would you find it useful to be able to access support programs using the Internet to assist you with your health problems? (n=111)** | | |
| Yes | | 82 (73.9) |
| No | | 29 (26.1) |
| **Would you find access through Internet useful, "yes", by age group, years (n=106)** | | |
| Under 50 (n=40) | | 25 (62.5) |
| 50 – 65 (n=54) | | 45 (83.3) |
| 65 and over (n=12) | | 8 (66.7) |
| **Would you find access through Internet useful, "yes", by socioeconomic area [b] (n=111)** | | |
| Low socioeconomic area (n=28) | | 21 (75.0) |
| Middle socioeconomic area (n=50) | | 39 (78.0) |
| High socioeconomic area (n=33) | | 22 (66.7) |
| **Reported reasons (n=71)** | | |
| Better understanding of health information and condition management | | 16 (22.5) |
| Avoid difficult travel requirements and less time consuming | | 13 (18.3) |
| Convenient and accessible help and information | | 12 (16.9) |
| "The more help the better" | | 9 (12.7) |
| Peer support and sharing | | 9 (12.7) |
| Address difficulty in accessing disease information from health providers | | 6 (8.5) |
| Up-to-date advice on management and treatments | | 6 (8.5) |
| **Are there health education or social group sites on the Internet that you have found helpful? (n=112)** | | |
| Yes | | 60 (53.6) |
| No | | 52 (46.4) |
| **Health education or social group sites helpful, "yes", by age group, years (n=107)** | | |
| Under 50 (n=38) | | 20 (52.6) |
| 50 – 65 (n=57) | | 29 (50.9) |
| 65 and over (n=12) | | 7 (58.3) |
| **Health education or social group sites helpful, "yes", by socioeconomic area [b] (n=112)** | | |
| Low socioeconomic area (n=30) | | 14 (46.7) |
| Middle socioeconomic area (n=48) | | 27 (56.3) |
| High socioeconomic area (n=34) | | 19 (55.9) |
| **Reported health education or social group sites (n=52)** | | |
| Health organization or research sites | | 16 (31) |
| | Australian Heart/Lung Transplant Association | 7 (43.8) |
| | Diabetes | 3 (18.8) |
| | Heart Lung Transplant Network | 1 (6.3) |
| | Arthritis Australia | 1 (6.3) |
| | Cystic Fibrosis | 1 (6.3) |
| | Hemochromatosis organization | 1 (6.3) |
| | Heart and lung sites | 1 (6.3) |
| | Heart foundation | 1 (6.3) |
| General browsing for health information and education | | 13 (25.0) |

| Online access, concerns and currently accessed sites | n (%)[a] |
|---|---|
| Medication and treatment information and side effects | 9 (17.9) |
| Facebook for disease-specific support groups | 8 (15.4) |
| Donate Life | 1 (1.9) |
| Health rebate and concession information | 1 (1.9) |
| Online mental health programs (Sadness and Depression program) | 1 (1.9) |
| **Would you have any concerns about participating in support programs via the Internet? (n=111)** | |
| Yes | 24 (21.6) |
| No | 87 (78.4) |
| **Concern about participating, "yes", by age group, years (n=106)** | |
| Under 50 (n=41) | 8 (19.5) |
| 50 – 65 (n=53) | 12 (22.6) |
| 65 and over (n=12) | 3 (25.0) |
| **Would you find access through Internet useful, "yes", by socioeconomic area [b] (n=111)** | |
| Low socioeconomic area (n=28) | 5 (17.9) |
| Middle socioeconomic area (n=50) | 13 (26.0) |
| High socioeconomic area (n=33) | 6 (18.2) |
| **Reported concerns (n=15)** | |
| Privacy and security | 10 (66.7) |
| Accuracy of information | 2 (13.3) |
| Computer literacy and access | 2 (13.3) |
| Limited Australian-based sites | 1 (6.7) |
| Misinterpretation of information | 1 (6.7) |
| No support group for my condition | 1 (6.7) |

[a] Multiple responses to questions were accepted in free text questions and respondents were instructed to "tick all that apply" when responding to nominal questions. In this context, the sum of percentages will be more than 100%.

[b] Australian Bureau of Statistics. Socioeconomic indexes for areas: robustness, diversity within larger areas, and the new geography standard Commonwealth of Australia 2012; ABS Catalogue no. 1351.0.55.038.

### Health Information and Education Websites Currently Accessed

The majority of respondents answered questions relating to health information and education sites currently accessed through the Internet (112/123, 91.1%). Over half of the respondents were already accessing websites that they felt were useful (60/112, 53.6%) and this was marginally higher in those aged 65 years and above (7/12, 58%) than those from middle (29/57, 51%) and low socioeconomic areas (20/38, 53%). A total of 52 (46.4%) of the 112 respondents reported commonly accessed sites, with health organizations and research sites (16/52, 31%), including Australian Heart/Lung Transplant Association, most common. One-third (16/52, 31%) of respondents stated that they did not access a particular website, but that they generally browsed the Internet for health information and education, with a further nine (17%) of 52 respondents accessing sites for medication and treatment information specifically. Eight (15%) of 52 responded that they accessed disease-specific Facebook support groups, and single individuals noted they accessed Donate Life, health rebate and concession sites, and an online mental health support program run by the academic medical center itself (1/52, 2%, respectively) (Table 3).

### Concerns Over Accessing Information and Support Online

When asked if respondents had concerns over accessing and participating in online support programs, the overwhelming majority of respondents answered the question (111/123, 90.2%) and did not have concerns (87/111, 78.4% answered "no"). This did not differ across gender, age, or socioeconomic groups. Reasons for concern were given by a small number of respondents (15/111, 13.5%), with privacy and security most common (10/15). Other reasons for individual concern included accuracy of information (2/15), computer literacy and access (2/15), limited Australian-based sites (1/15), the opportunity for misinterpretation of information (1/15), and the lack of a support group for that individual's particular condition (1/15) (Table 3).

## Discussion

### Principal Findings

Web-based health information and support are available in the United States [28, 32-35]; however, internationally, the use is not as widespread [28,30-36]. In the Australian context, studies have explored Web-based health interventions, but there is limited information as to the patterns of technology use in cardiopulmonary patients [37]. The reported study sought to describe patterns of technology use in patients attending a cardiopulmonary clinic. The patients responding to this survey provide a sample of those individuals likely to have increasing health care utilization needs. It is important to take this initial step in understanding whether these consumers have technology capabilities and receptivity to these modalities [42-44].

Study results indicated that computer, mobile phone, and Internet use are a pervasive part of everyday life with individuals using their technological devices for a variety of reasons, including accessing and browsing health information websites. The majority of respondents additionally answered that access to support programs through a telecommunication platform would provide assistance with their health problems; this was most common in individuals aged 50-65 years. The most commonly accessed websites were disease-specific sites, organizations, and research sites, as well as sites that provided information on specific medications and treatments. In agreement with previous literature, peer support and sharing of experiences were also noted as benefits of access through an online platform, and was noted as providing support and information that they may not otherwise be able to access in their everyday life [29-31,46,47].

Patterns of technology use did differ between patients from different socioeconomic groups, as measured using advantage/disadvantage index based on area [45]. Although overall technology use was pervasive in all groups, patients who lived in higher socioeconomic areas used the Internet most regularly and those patients from middle socioeconomic areas were most likely to access social media compared with the other groups. Although still high users of technology, those from low socioeconomic areas had less access to computers and used computers and the Internet less frequently, a situation evident in international literature; technological access and literacy are a consideration for future technology-based health delivery interventions [48,49]. Interestingly, while those who lived in higher socioeconomic areas were most likely to be already accessing Web-based health information sites, when asked if they would benefit from delivery of health information and support through an online interface, those from higher socioeconomic areas were least likely to respond positively; this may reflect higher health literacy [27,50], better health access [2,45], and therefore less need for additional support, but this would need further investigation.

Respondents indicated clear issues with current health care delivery through face-to-face interaction, with several noting the long distance they had to travel to access care and the lack of information provision through traditional sources; Web-based health information delivery may go some way to alleviating the limitations of current health care provision. As similarly noted in previous literature, respondents viewed online health care delivery as providing convenient, timely, and easily accessible information, currently difficult to obtain through traditional face-to-face sources [29-31,51]. This is particularly important in the context of this cardiopulmonary clinic, which acts as a quaternary referral clinic attracting patients from a large geographical area across the state. Several studies have highlighted the relationship between patient satisfaction and Web-based health information seeking behavior [51-53]. Consumer-health provider interfaces need to be improved to provide timely and accessible health care interaction that reduces the geographical burden of current health care delivery [28-31,51-53].

While the majority of respondents stated that they did not have concerns over accessing information or support online, issues of privacy and security, the accuracy of information, and the potential for misinterpretation of information were raised by a smaller number of patients. Consumers' ability to distinguish accurate, trustworthy, and personally applicable information, when faced with the sheer volume of health information sites available, is a commonly reported challenge in the literature [42-44]. Development and validation of websites is essential; health professionals have an opportunity to ensure that patients and their families have guidance to accurate and trustworthy Web-based health information sources [42-44].

Web-based health care delivery has particular potential to provide convenient and accessible access for individuals and their families living with chronic, complex, and progressive conditions [28-31]. Providing ongoing care through technology platforms may address the issues associated with short-term episodic programs, such as pulmonary and cardiac rehabilitation, in providing ongoing education, social support, and exercise maintenance to larger patient cohorts [16,17,21,22,28]. Self-management programs that are provided through a Web-based interface may leverage computer-based and mobile technology to facilitate continued care and support [28-31]. This may be of particular value to aging "baby boomers", who have already incorporated these technologies into their daily lives [28-31]. Web-based health care delivery additionally has the potential to help those at end of life who need increasingly complex strategies to cope with dyspnea and fatigue, especially as they become homebound [28-31].

### Implications for Practice

This study sought to describe patterns of technology use in patients attending a cardiopulmonary clinic. Technology use is a pervasive part of everyday life regardless of age or socioeconomic group with patients already heavily engaged in health-seeking behaviors through Web-based sources. There is a necessity to develop and validate websites, and an opportunity to ensure that patients and their families have guidance to accurate and credible health information sources. Web-based delivery of health information and support is of particular importance in patients with cardiopulmonary disease, who are most likely to have high symptom burden and health care utilization needs in the future. Current consumer-health provider interfaces need to be improved to accommodate the changing needs and preferences of an empowered generation of health consumers, and to provide timely and accessible health care interaction that addresses the geographical burden of current health care delivery. Telecommunication interfaces may alleviate some of the difficulties with current health care access and provide an increasingly important adjunct to traditional forms of health care delivery. We are at a turning point within the evolution of health care delivery and have the opportunity to shape how future interventions deliver health information and promote self-management. Future research must explore the feasibility of delivering health care through Web-based platforms across larger cohorts and explore the social and economic impact of this approach on health care delivery.

## Limitations

This prevalence study was undertaken in a small cohort of patients from a single clinical site. While survey responses were completed in full in most cases, there is a possibility that patients who do not engage with technology may have self-excluded from participating. Further, large cohort, multi-site research would be required to describe overall population technology use. Additionally, this study only sought to describe the patterns of technology use and further research is required to understand the attitudes and specific barriers faced by cardiopulmonary pulmonary patients in regard to the delivery of health information and education through telecommunication interfaces. While this initial study does have its limitations, the results do provide important information regarding patients' access to technological devices, their use of Web-based information and support for their health conditions, and the perceived potential benefits of health care delivery through Web-based platforms. This is particularly important in the context of patients attending a cardiopulmonary clinic, who are most likely to have high symptom burden and associated health care utilization needs in the future.

## Conclusions

Chronic disease burden and the long-term self-management tasks that challenge patients with cardiopulmonary disease are a compelling argument for accessible and convenient avenues to obtaining ongoing treatment and peer support. Technology use was already a pervasive part of everyday life for study participants, and a central platform for health care interactions including common access of health information and education. Patterns of use and access differed marginally across age and socioeconomic groups, however, accessing Web-based health information was prevalent for all groups. Clear issues were raised over long distance travel and a lack of information provision through traditional health delivery sources. Web-based access to health support programs are perceived as providing convenient, timely, and easily accessible information—particularly important in the context of a quaternary referral clinic attracting patients from a large geographical area, and in cardiopulmonary patients most likely to have high health care utilization needs in the future. Telecommunication interfaces will be an increasingly important adjunct to traditional forms of health care delivery. These will need to be assessed for the validity of content and access to target populations.

---

## Acknowledgments

We would like to acknowledge the patients and staff at St Vincent's Hospital, Sydney, Australia for their participation and assistance in this study.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Internet access and use survey.

[PDF File (Adobe PDF File), 6KB - ijmr_v4i1e5_app1.pdf ]

## References

1. Beaglehole R, Bonita R, Alleyne G, Horton R, Li L, Lincoln P, Lancet NCD Action Group. UN high-level meeting on non-communicable diseases: addressing four questions. Lancet 2011 Jul 30;378(9789):449-455. [doi: 10.1016/S0140-6736(11)60879-9] [Medline: 21665266]
2. Australia's Health 2008, Cat. no. 99. Canberra: Australian Institute of Health and Welfare; 2008. URL: http://www.aihw.gov.au/publication-detail/?id=6442468102 [accessed 2015-02-13] [WebCite Cache ID 6WJixRbdc]
3. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. Nat Rev Cardiol 2011 Jan;8(1):30-41 [FREE Full text] [doi: 10.1038/nrcardio.2010.165] [Medline: 21060326]
4. McMurray JJ, Petrie MC, Murdoch DR, Davie AP. Clinical epidemiology of heart failure: public and private health burden. Eur Heart J 1998 Dec;19 Suppl P:P9-16. [Medline: 9886707]
5. Mahmood SS, Wang TJ. The epidemiology of congestive heart failure: the Framingham Heart Study perspective. Glob Heart 2013 Mar 1;8(1):77-82 [FREE Full text] [doi: 10.1016/j.gheart.2012.12.006] [Medline: 23998000]
6. Mannino DM, Buist AS. Global burden of COPD: risk factors, prevalence, and future trends. Lancet 2007 Sep 1;370(9589):765-773. [doi: 10.1016/S0140-6736(07)61380-4] [Medline: 17765526]

7. Fabbri L, Pauwels RA, Hurd SS, GOLD Scientific Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD Executive Summary updated 2003. COPD 2004 Apr;1(1):105-41; discussion 103. [doi: 10.1081/COPD-120030163] [Medline: 16997745]
8. Blinderman CD, Homel P, Billings JA, Tennstedt S, Portenoy RK. Symptom distress and quality of life in patients with advanced chronic obstructive pulmonary disease. J Pain Symptom Manage 2009 Jul;38(1):115-123. [doi: 10.1016/j.jpainsymman.2008.07.006] [Medline: 19232893]
9. Currow DC, Smith J, Davidson PM, Newton PJ, Agar MR, Abernethy AP. Do the trajectories of dyspnea differ in prevalence and intensity by diagnosis at the end of life? A consecutive cohort study. J Pain Symptom Manage 2010 Apr;39(4):680-690. [doi: 10.1016/j.jpainsymman.2009.09.017] [Medline: 20413056]
10. Edmonds P, Karlsen S, Khan S, Addington-Hall J. A comparison of the palliative care needs of patients dying from chronic respiratory diseases and lung cancer. Palliat Med 2001 Jul;15(4):287-295. [Medline: 12054146]
11. Elkington H, White P, Addington-Hall J, Higgs R, Edmonds P. The healthcare needs of chronic obstructive pulmonary disease patients in the last year of life. Palliat Med 2005 Sep;19(6):485-491. [Medline: 16218161]
12. Skilbeck J, Mott L, Page H, Smith D, Hjelmeland-Ahmedzai S, Clark D. Palliative care in chronic obstructive airways disease: a needs assessment. Palliat Med 1998 Jul;12(4):245-254. [Medline: 9743823]
13. Fletcher MJ, Upton J, Taylor-Fishwick J, Buist SA, Jenkins C, Hutton J, et al. COPD uncovered: an international survey on the impact of chronic obstructive pulmonary disease [COPD] on a working age population. BMC Public Health 2011;11:612 [FREE Full text] [doi: 10.1186/1471-2458-11-612] [Medline: 21806798]
14. Corsonello A, Antonelli Incalzi R, Pistelli R, Pedone C, Bustacchini S, Lattanzio F. Comorbidities of chronic obstructive pulmonary disease. Curr Opin Pulm Med 2011 Dec;17 Suppl 1:S21-S28. [doi: 10.1097/01.mcp.0000410744.75216.d0] [Medline: 22209926]
15. Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. JAMA 2002 Nov 20;288(19):2469-2475. [Medline: 12435261]
16. Cecins N, Geelhoed E, Jenkins S. Reduction in hospitalisation following pulmonary rehabilitation in patients with COPD. Aust Health Rev 2008 Aug;32(3):415-422. [Medline: 18666869]
17. Nici L, Donner C, Wouters E, Zuwallack R, Ambrosino N, Bourbeau J, ATS/ERS Pulmonary Rehabilitation Writing Committee. American Thoracic Society/European Respiratory Society statement on pulmonary rehabilitation. Am J Respir Crit Care Med 2006 Jun 15;173(12):1390-1413. [doi: 10.1164/rccm.200508-1211ST] [Medline: 16760357]
18. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the Patient Care Committee of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2011 Feb;13(2):115-126. [doi: 10.1093/eurjhf/hfq219] [Medline: 21148593]
19. Clark AM, Davidson P, Currie K, Karimi M, Duncan AS, Thompson DR. Understanding and promoting effective self-care during heart failure. Curr Treat Options Cardiovasc Med 2010 Jan;12(1):1-9. [doi: 10.1007/s11936-009-0053-1] [Medline: 20842477]
20. Krum H, Jelinek MV, Stewart S, Sindone A, Atherton JJ, National Heart Foundation of Australia, Cardiac Society of Australia and New Zealand. 2011 update to National Heart Foundation of Australia and Cardiac Society of Australia and New Zealand Guidelines for the prevention, detection and management of chronic heart failure in Australia, 2006. Med J Aust 2011 Apr 18;194(8):405-409. [Medline: 21495941]
21. RWJ Foundation. Patient education and consumer activation in chronic disease.: The Robert Wood Johnson Foundation; 2000. URL: http://www.rwjf.org/publications/publicationsPdfs/patient_education.pdf [accessed 2015-01-05] [WebCite Cache ID 6VLdttH5b]
22. Connors GL, Hodgkin JE, Celli BR. Pulmonary rehabilitation: guidelines to success. Philadelphia: Lippincott Williams & Wilkins; 2000.
23. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [Medline: 12867348]
24. Goodman RA, Parekh AK, Koh HK. Toward a more cogent approach to the challenges of multimorbidity. Ann Fam Med 2012;10(2):100-101 [FREE Full text] [doi: 10.1370/afm.1391] [Medline: 22412000]
25. Erbas B, Ullah S, Hyndman RJ, Scollo M, Abramson M. Forecasts of COPD mortality in Australia: 2006-2025. BMC Med Res Methodol 2012;12:17 [FREE Full text] [doi: 10.1186/1471-2288-12-17] [Medline: 22353210]
26. Kruis AL, Smidt NW, Assendelft W, Gussekloo J, Boland M, Rutten-van Mölken M, et al. Integrated disease management interventions for patients with chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2013;10:CD009437. [doi: 10.1002/14651858.CD009437.pub2] [Medline: 24108523]
27. Taggart J, Williams A, Dennis S, Newall A, Shortus T, Zwar N, et al. A systematic review of interventions in primary care to improve health literacy for chronic disease behavioral risk factors. BMC Fam Pract 2012;13:49 [FREE Full text] [doi: 10.1186/1471-2296-13-49] [Medline: 22656188]
28. Zickuhr K, Madden M. Pew Internet & American Life Project. Washington, DC: Pew Research Center’s Internet & American Life Project; 2012. Older adults and internet use URL: http://www.pewinternet.org/2012/06/06/older-adults-and-internet-use/ [accessed 2015-01-05] [WebCite Cache ID 6VLeXySGn]

29. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69. [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]
30. Wicks P, Massagli M, Frost J, Brownstein C, Okun S, Vaughan T, et al. Sharing health data for better outcomes on PatientsLikeMe. J Med Internet Res 2010;12(2):e19 [FREE Full text] [doi: 10.2196/jmir.1549] [Medline: 20542858]
31. Mo PK, Coulson NS. Developing a model for online support group use, empowering processes and psychosocial outcomes for individuals living with HIV/AIDS. Psychol Health 2012;27(4):445-459. [doi: 10.1080/08870446.2011.592981] [Medline: 21854088]
32. Nguyen HQ, Carrieri-Kohlman V, Rankin SH, Slaughter R, Stulbarg MS. Is Internet-based support for dyspnea self-management in patients with chronic obstructive pulmonary disease possible? Results of a pilot study. Heart Lung 2005;34(1):51-62. [doi: 10.1016/j.hrtlng.2004.06.005] [Medline: 15647734]
33. Nguyen HQ, Donesky-Cuenco D, Wolpin S, Reinke LF, Benditt JO, Paul SM, et al. Randomized controlled trial of an internet-based versus face-to-face dyspnea self-management program for patients with chronic obstructive pulmonary disease: pilot study. J Med Internet Res 2008;10(2):e9 [FREE Full text] [doi: 10.2196/jmir.990] [Medline: 18417444]
34. Nguyen HQ, Gill DP, Wolpin S, Steele BG, Benditt JO. Pilot study of a cell phone-based exercise persistence intervention post-rehabilitation for COPD. Int J Chron Obstruct Pulmon Dis 2009;4:301-313 [FREE Full text] [Medline: 19750190]
35. Nguyen HQ, Donesky D, Reinke LF, Wolpin S, Chyall L, Benditt JO, et al. Internet-based dyspnea self-management support for patients with chronic obstructive pulmonary disease. J Pain Symptom Manage 2013 Jul;46(1):43-55. [doi: 10.1016/j.jpainsymman.2012.06.015] [Medline: 23073395]
36. Murray E, Burns J, See TS, Lai R, Nazareth I. Interactive health communication applications for people with chronic disease. Cochrane Database Syst Rev 2005(4):CD004274. [doi: 10.1002/14651858.CD004274.pub4] [Medline: 16235356]
37. Biezen R, Jones KM, Piterman L. Managing chronic disease--patients' views and attitudes to using a broadband based service. Aust Fam Physician 2013 Aug;42(8):578-581 [FREE Full text] [Medline: 23971069]
38. Kwankam SY. What e-Health can offer. Bull World Health Organ 2004 Oct;82(10):800-802 [FREE Full text] [Medline: 15643805]
39. World Health Organization. E-Health. 2014. URL: http://www.who.int/trade/glossary/story021/en/ [accessed 2015-01-05] [WebCite Cache ID 6VLemhMUw]
40. Kay M. mHealth: New horizons for health through mobile technologies.: World Health Organization; 2011. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2015-01-05] [WebCite Cache ID 6VLeucCzO]
41. de Jong CC, Ros WJ, Schrijvers G. The effects on health behavior and health outcomes of Internet-based asynchronous communication between health providers and patients with a chronic condition: a systematic review. J Med Internet Res 2014 Jan;16(1):e19 [FREE Full text] [doi: 10.2196/jmir.3000] [Medline: 24434570]
42. Adams SA. Revisiting the online health information reliability debate in the wake of "web 2.0": an inter-disciplinary literature and website review. Int J Med Inform 2010 Jun;79(6):391-400. [doi: 10.1016/j.ijmedinf.2010.01.006] [Medline: 20188623]
43. Fox S, Duggan M. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013. Health online 2013 URL: http://www.pewinternet.org/files/old-media/Files/Reports/PIP_HealthOnline.pdf [accessed 2015-01-05] [WebCite Cache ID 6VLf1jGGe]
44. Siliquini R, Ceruti M, Lovato E, Bert F, Bruno S, De Vito E, et al. Surfing the internet for health information: an Italian survey on use and population choices. BMC Med Inform Decis Mak 2011;11:21 [FREE Full text] [doi: 10.1186/1472-6947-11-21] [Medline: 21470435]
45. ABS Catalogue no. 1351.0.55.038. Canberra: Australian Bureau of Statistics; 2012 Mar. Socio-economic indexes for areas: Robustness, diversity within larger areas and the new geography standard Commonwealth of Australia URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/1351.0.55.038Mar%202012?OpenDocument [accessed 2015-02-13] [WebCite Cache ID 6WJlLOv44]
46. van der Eijk M, Faber MJ, Aarts JW, Kremer JA, Munneke M, Bloem BR. Using online health communities to deliver patient-centered care to people with chronic conditions. J Med Internet Res 2013 Jun;15(6):e115 [FREE Full text] [doi: 10.2196/jmir.2476] [Medline: 23803284]
47. Bartlett YK, Coulson NS. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. Patient Educ Couns 2011 Apr;83(1):113-119. [doi: 10.1016/j.pec.2010.05.029] [Medline: 20599338]
48. Wangberg SC, Andreassen HK, Prokosch HU, Santana SM, Sørensen T, Chronaki CE. Relations between Internet use, socio-economic status (SES), social support and subjective health. Health Promot Int 2008 Mar;23(1):70-77 [FREE Full text] [doi: 10.1093/heapro/dam039] [Medline: 18083686]
49. Wei L. Number matters: the multimodality of Internet use as an indicator of the digital inequalities. Journal of Computer‐Mediated Communication 2012;17(3):303-318. [doi: 10.1111/j.1083-6101.2012.01578.x]
50. Dennis S, Williams A, Taggart J, Newall A, Denney-Wilson E, Zwar N, et al. Which providers can bridge the health literacy gap in lifestyle risk factor modification education: a systematic review and narrative synthesis. BMC Fam Pract 2012;13:44 [FREE Full text] [doi: 10.1186/1471-2296-13-44] [Medline: 22639799]

51. 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]
52. Hou J, Shim M. The role of provider-patient communication and trust in online sources in Internet use for health-related activities. J Health Commun 2010;15 Suppl 3:186-199. [doi: 10.1080/10810730.2010.522691] [Medline: 21154093]
53. Higgins O, Sixsmith J, Barry M, Domegan C. A literature review on health information seeking behaviour on the web: a health consumer and health professional perspective.: European Centre for Disease Prevention and Control; 2011. URL: http://www.ecdc.europa.eu/en/publications/Publications/Literature%20review%20on%20health%20information-seeking%20behaviour%20on%20the%20web.pdf [accessed 2015-01-05] [WebCite Cache ID 6VLfbXMNf]

---

## Abbreviations

**CHF:** congestive heart failure
**COPD:** chronic obstructive pulmonary disease

---

*Edited by G Eysenbach; submitted 20.10.14; peer-reviewed by H Du, R McNamara, A Lau; comments to author 12.11.14; revised version received 30.11.14; accepted 20.12.14; published 06.03.15.*

*Please cite as:*
*Disler RT, Inglis SC, Newton PJ, Currow DC, Macdonald PS, Glanville AR, Donesky D, Carrieri-Kohlman V, Davidson PM*
*Patterns of Technology Use in Patients Attending a Cardiopulmonary Outpatient Clinic: A Self-Report Survey*
*Interact J Med Res 2015;4(1):e5*
*URL: http://www.i-jmr.org/2015/1/e5/*
*doi:10.2196/ijmr.3955*
*PMID:25798814*

Original Paper

# Understanding Internet Use Among Dementia Caregivers: Results of Secondary Data Analysis Using the US Caregiver Survey Data

Heejung Kim[1], RN, PhD

University of Kansas, School of Nursing, Kasnas City, KS, United States

**Corresponding Author:**
Heejung Kim, RN, PhD
University of Kansas
School of Nursing
MS 4043
3901 Rainbow Blvd
Kasnas City, KS, 66160
United States
Phone: 1 913 588 1657 ext 8
Fax: 1 913 588 1660
Email: hkim2@kumc.edu

## Abstract

**Background:** Informal caregivers of persons with dementia experience higher levels of chronic stress in the caregiving trajectory. The Internet provides diverse types of caregiver resources that may help ameliorate their stress and relevant negative outcomes. However, there is limited information about the prevalence and factors of using Internet-based resources for health- and caregiving-related purposes in informal caregivers of persons with dementia.

**Objective:** Specific aims of this study were (1) to determine the prevalence and factors of caregiver's health-related Internet use and (2) to compare sociodemographic and caregiving-related characteristics between health-related Internet users and non–health-related Internet users among informal caregivers of persons with dementia.

**Methods:** This quantitative investigation was a descriptive correlational design using a secondary data analysis. Primary data were collected via a survey conducted in 2009 by the National Alliance for Caregiving and the American Association of Retired Persons. Telephone interviews utilizing standardized questionnaires were used to collect self-reported information about sociodemographics and caregiving-related history (N=450). Descriptive statistics and a hierarchical binary logistic regression analysis were completed based on the stress process model.

**Results:** Approximately 59% (265/450) of dementia caregivers were identified as health-related Internet users. Caregivers' sociodemographics and their subjective responses of caregiving stress were the most significant factors to identify health-related Internet users followed by workload assisting in instrumental activities of daily living of persons with dementia. There were significant differences for caregiver's age, levels of education and income, hours spent caregiving, and the relationship to persons with dementia between health-related Internet users and non–health-related Internet users ($P<.05$ for all). After controlling for confounding effects, younger age of persons with dementia (OR 0.278, 95% CI 0.085-0.906), higher education levels of caregivers (OR 3.348, 95% CI 2.019-5.552), shorter caregiving time spent per week (OR 0.452, 95% CI 0.243-0.840), higher levels of caregiver's emotional stress (OR 1.249, 95% CI 1.004-1.555), and financial hardship (OR 4.61, 95% CI 1.416-14.978) were identified as newly emerging factors of health-related Internet use.

**Conclusions:** Although the Internet provided useful resources for caregivers of persons with dementia, dementia caregivers reported lower levels of health-related Internet use compared to the general public. Our findings confirmed the impact of age, education levels, and/or income on Internet use reported in previous studies. However, the predictive value of subjective responses of caregiving stress for health-related Internet use was a new addition. These findings will assist health care providers, researchers, and policy makers in identifying who is the least likely to access Internet-based resources and how Internet-based strategies can best be designed, implemented, and distributed to meet the needs of this group of users.

***(Interact J Med Res 2015;4(1):e1)*** doi:10.2196/ijmr.3127

**KEYWORDS**

Internet; dementia; caregiver; stress; consumer health information

## Introduction

Dementia, including Alzheimer disease, refers to cognitive disorders presenting memory impairment, difficulty in language, organizational ability, abstract thinking, object recognition, and disturbance of executive function [1]. Adults caring for persons living with dementia (hereafter dementia caregivers) are the second-largest informal caregiver group in the United States because most people with dementia are older adults aged 65 or older [1,2]. Dementia caregivers are more likely to experience a wide range of negative behaviors or health problems than persons with nondementia health problems [2,3]. For example, dementia caregivers frequently exhibit maladaptive coping strategies, express concern about their poor quality of life, experience lower self-rated health, and report a higher level of caregiver burden [4-6]. In addition, dementia caregivers report severe sleep disturbances, clinical depression, and higher mortality compared to other caregivers [4,6-8]. Interestingly, these negative effects of dementia caregiving project to their care recipients because dyads of caregivers and persons with dementia are interdependent in the family unit [9]. Caregiver stress and burden have been shown to increase caregiver's harmful or abusive behaviors toward their care recipients [10], accelerate the early placement of persons with dementia into institutional care [11], and decrease the life expectancy of care recipients [12]. Thus, timely reduction of caregiver stress and related problems are critical for both caregivers and care recipients.

To discontinue this vicious cycle between caregivers and care recipients, the stress process model emphasizes the proper use of resources to mediate the relationship between caregiver stress and relevant consequences [13]. Caregivers will experience higher levels of stress if they perceive their demands to be beyond the capacity of their coping resources [13]. Dementia caregivers are likely to seek out external resources that will help them resolve their stress, manage their health problems, and provide ongoing care for persons with dementia [2,3,12]. However, previous studies of resource use have generally focused on traditional face-to-face resources, such as professional health care services and support [6]; community-based services, such as respite services or caregiving assistance from professionals or nonprofessionals [14]; or agency-provided health and human resources [15]. However, there are no investigations examining Internet-based resources for health-related purposes in dementia caregivers.

Internet-based health resources include health information on websites and activities via communication technologies, whereas it excludes specific health interventions based on information and communication technology (ICT) specifically designed by clinical researchers [16,17]. Internet-based resources are now widely used and well integrated into the daily lives of the caregiving population [18]. The Internet modality assists in overcoming the limitations of a face-to-face approach; namely, time constraints, geographic limitations, and transportation issues [19,20]. Surveys of family caregivers report that 80% to 95% request technology-based interventions and Internet-based information or resources for enhancing better caregiving on behalf of their care recipients [21,22]. Clinicians have suggested various online information to caregivers as practical adjuncts or alternatives to traditional approaches without proven evidence [23]. However, it has rarely been evaluated in terms of how much dementia caregivers use the Internet for health-related purposes (hereafter health-related Internet use) and what factors affect their use. Thus, evidence-based practice requires more data to support current clinical practice.

This study used a modified stress process model by adding health-related Internet use as new subcomponent of resources. The original framework, developed by Pearlin et al [13] in 1990, has been used to understand how caring for a person with dementia affects both the health and well-being of both persons with dementia and their informal caregivers. In the multidimensional caregiving stress process, appropriate use of external resources mediates their coping within stressful situations. The addition of health-related Internet use to this original model can reflect the current need for virtual care resources for dementia caregivers living in the high-tech society of the 21st century [21,22]. The modified model was used for this study regarding (1) defining study constructs, (2) selecting and operationalizing study variables, (3) guiding data analysis, and (4) interpreting findings for clinical inferences.

This study evaluated how dementia caregivers use Internet-based resources for health- and caregiving-related purposes. Specifically, this study examined the prevalence and factors of health-related Internet use in dementia caregivers as well as differences between Internet users and non-Internet users. Three research questions were proposed:

1. What percentage of dementia caregivers use Internet-based resources for health- and caregiving-related purposes?
2. Are there any differences in sociodemographic and caregiving-related characteristics between health-related Internet users and non–health-related Internet users?
3. Which sociodemographic and caregiving-related factors are associated with health-related Internet use in dementia caregivers?

## Methods

### Design

The study was a cross-sectional and descriptive correlational design using a secondary data analysis. The primary data source was from the National Alliance for Caregiving (NAC) and the American Association of Retired Persons (AARP). This dataset was selected for this study because of (1) its up-to-date information on Internet use by dementia caregivers and (2) continuous refinement of sampling and data collection over the past decade [24,25].

### Description of the Primary Data Source

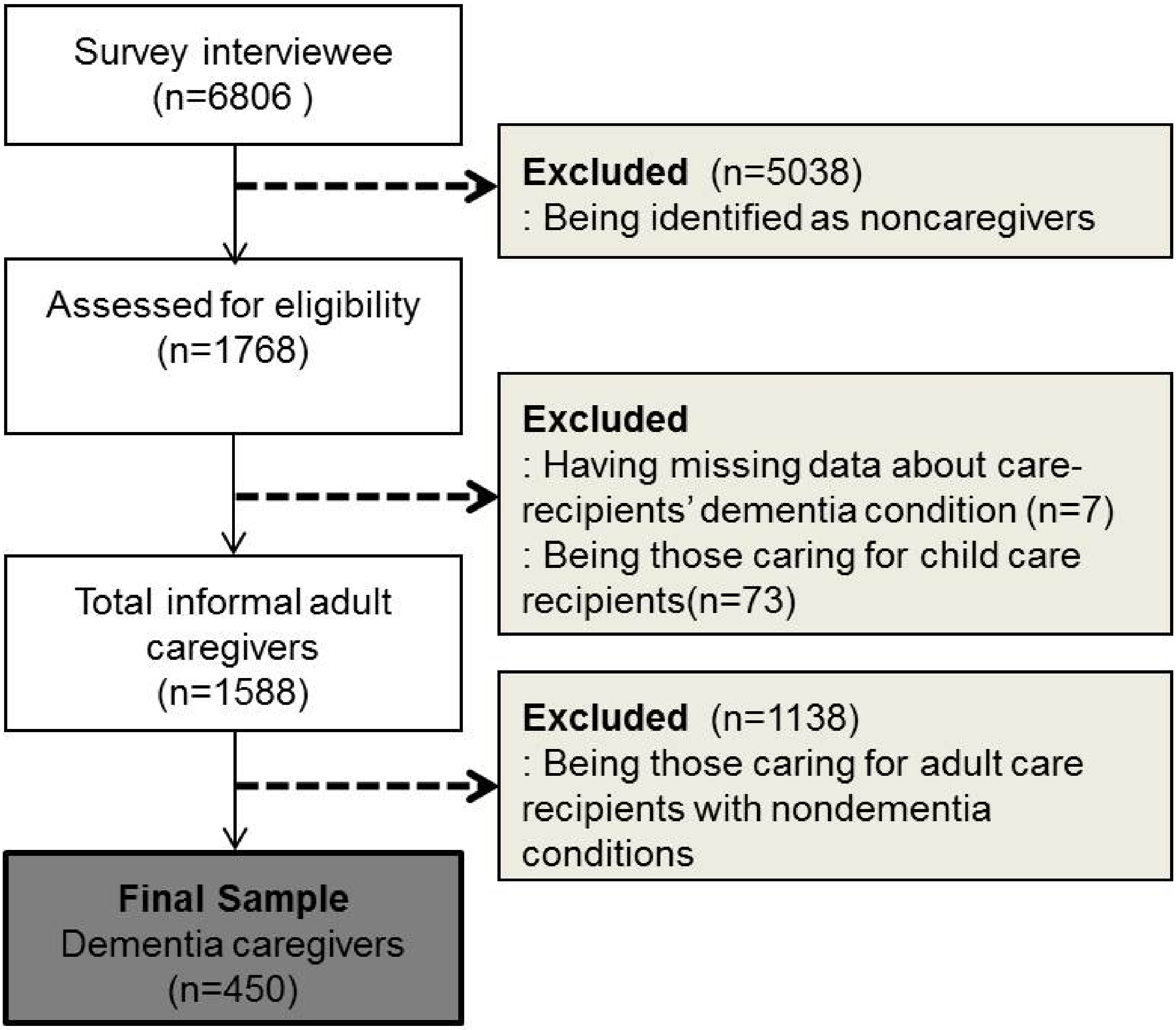
The NAC and AARP survey collected sociodemographic and caregiving-related data about persons with dementia and caregivers as well as their Internet and technology use. Interviews using a standardized questionnaire were programmed into a computerized telephone system and were conducted from March to June 2009. Interviewees were 6806 adults living in communities in California, Delaware, Illinois, Kansas, Ohio, Virginia, and the state of Washington in the United States.

Random digit dialing based on surnames produced a set of telephone numbers stratified by geographic population density. Oversampling was done in racial and ethnic minorities (African American, Asian, and Hispanic groups) and older adults (age 50 years or older). One respondent was randomly selected from each household. If there were multiple care recipients (2 or more) for 1 caregiver, interviewers focused on the information for 1 primary care recipient who was receiving the most assistance from the caregiver [24,25].

## Sample

This study used a subset of the data belonging to persons with dementia and their caregivers. Dementia caregivers were defined as persons who provided unpaid care or assistance to a family, relative, friend, or anyone living with Alzheimer dementia, any other type of dementia, or dementia-related conditions (confusion or forgetfulness). Among the 1768 informal caregivers in the dataset, 450 eligible caregivers were included in this study after excluding those caring for persons younger than 18 years (n=173), anyone with incomplete data regarding their care recipients' dementia condition (n=7), and nondementia caregivers (n=1138) (See Figure 1).

**Figure 1.** Flowchart of study samples.



## Measures

### *Baseline Information on Dementia Care Recipients and Their Caregivers*

Both persons with dementia and their caregiver's sociodemographic information were collected: age, gender, race/ethnicity, education level, residence area, and income, as well as the relationship between persons with dementia and their caregivers. Ages were a continuous variable, whereas all others were categorical variables. References (coded as 0) were those who were male, non-Hispanic white, had less than college-level education, rural residents, nonfamily or nonrelatives, and a household income of less than US $30,000 per year.

### *Caregiving-Related Information*

Dementia caregivers self-reported the number of hours of caregiving per week, the duration of caregiving in years, and subjective responses of caregiving stress. The number of hours spent on caregiving tasks indicated how many hours they devoted to caregiving per week (range 1-168): 1 indicated they spent 1 hour or less per week and 168 indicated they engaged in full-time caregiving work. The duration of the caregiving indicated how long they had been performing the caregiver role. Here, 1 indicated that at the time of the survey they had either spent 1 year or less as a caregiver or that they only occasionally provided caregiving on an on-and-off basis. Higher values represented the approximate number of years they had been providing care.

Objective caregiving stressors were primarily based on a functional dependency of persons with dementia in terms of activities of daily living (ADLs) and instrumental activities of daily living (IADLs). Here, the ADLs consisted of 6 activities: transferring, dressing, toileting, bathing, feeding, and handling incontinence or diapers (range 0-6) [26]. The IADLs were selected based on Lawton and Brody’s scale (range 0-6) [27] and included 6 activities: managing medications, managing finances, shopping, doing housework including laundry, preparing meals, and transportation. Higher scores indicated that persons with dementia were more dependent on caregivers in their daily living and for instrumental functions. In this study sample, Cronbach alpha of ADLs was .82 and the IADLs was .70. Subjective responses of caregiving stress were evaluated in terms of physical strain, emotional stress, and financial hardship. Each item was scored on a 5-point Likert scale, with higher scores indicating more physical strain, emotional stress, and financial hardship experienced. Moderate correlations were observed across the 3 items.

### *Health-Related Internet Use*

The main focus of interest for this study was health-related Internet use, measured by self-reports of the frequency of Internet use for health-related purposes. Frequency of health-related Internet use was measured by asking the question, “How often, if at all, have you gone to Internet websites in the past year to find information and resources in any way related to being a caregiver for your care recipient? Often, sometimes, rarely, or never?” [16] This study used a consistent definition of health-related Internet use, which has been used in previous studies to compare prevalence [16,17]. Non-health-related Internet users were defined as those who never used Internet-based resources for health and caregiving purposes. All others were defined as health-related Internet users.

### Procedures

The data acquisition and use was approved by the NAC and AARP. All data provided were deidentified to follow Health Insurance Portability and Accountability Act privacy rules. The University of Virginia’s Institutional Review Board for Social and Behavioral Sciences reviewed this project and confirmed the exempt status. Data analyses included both preliminary analyses (reliability tests, intercorrelation analyses, an exploratory factor analyses, and relevant statistical assumptions checks) and main analyses (descriptive analyses and a hierarchical binary logistic regression analysis). A total of 13 cases were dropped from the main analysis based on listwise deletion across all independent and dependent variables. This represented only 2.9% of the sample; thus, no data imputation was conducted [28].

Before the descriptive and regression analysis, all statistical assumptions were checked including univariate/multivariate normality, linearity, and multicollinearity. To correct for univariate normality, the variable of age of persons with dementia, number of hours for caregiving, and financial hardship was transformed using a $\log_{10}$ function. For multivariate normality, 2 outliers were identified based on the Mahalanobis distance function. When comparing results from a model with 2 outliers to those from the model without the outliers, there were no differences in $R^2$, coefficients, *F* statistics, or *P* values, although the values of the Mahalanobis distance were corrected. To glean the maximum information from the available samples, the final results were reported from the model that did not exclude those 2 outliers. The final sample size (N=437) in the main analysis was sufficient for conducting multiple regression with 16 independent variables because the suggested sample size was 160-320 [28-30].

### Statistical Analysis

To answer the first and second research questions, percentage responding weighted frequency and means (SD) were reported as well as results of univariate descriptive statistics (independent *t* tests, Mann-Whitney *U* tests, or chi-square tests) to compare health-related Internet users to non-health-related Internet users in the dementia caregivers. To answer the third research question, a hierarchical binary logistic regression analysis was completed to identify factors of the health-related Internet user group (0=non-health-related Internet users; 1=health-related Internet users). Based on the stress process model [13], 16 independent variables were entered into the regression model. Block 1 included age and gender of persons with dementia, whereas Block 2 included sociodemographics of caregivers, including age, gender, education level, income, race and ethnicity, resident care, and relationship to persons with dementia. Separately, functional dependency in terms of ADLs and IADLs was included in Block 3 and caregiving history was included in Block 4. Subjective responses of caregiving stress were included in Block 5. SPSS 20.0 (IBM Corp, Armonk, NY, USA) was used for data analyses. The significance-level criterion for all statistical tests was alpha=.05, 2-tailed. To infer generalizable findings, the study applied a composite score for the population weight, which was calculated based on age, gender, and race/ethnicity, and the results compared to the 2008 population estimates released by the Population Division of the US Census Bureau on May 14, 2009 [24].

## *Results*

### Sample Characteristics of Persons With Dementia and Their Caregivers

The mean age of persons with dementia was 78.37 years (SD 14.13) and the majority were women (69.1%, 309/447). They had moderate levels of functional impairments of ADLs and IADLs (mean 2.21, SD 2.01 and mean 4.21, SD 1.71, respectively). The mean age of their caregivers was 50.30 years (SD 14.98). The majority of them were women (62.0%, 277/447), non-Hispanic whites (71.6%, 320/450), and children or grandchildren of persons with dementia (74.9%, 335/447). Caregivers were educated at the level of high school or less (53.5%, 240/449) and had overall household incomes greater than US $30,000 per year (70.3%, 314/447). The residence areas of the caregivers were evenly distributed (urban: 29.3%, 131/447; suburban: 37.8%, 169/447; rural: 30.7%, 137/447). Caregivers spent a mean 29.96 (SD 46.93) hours per week for a mean of 5.22 (SD 7.96) years performing the role of caregiver for the person with dementia.

### Comparison of Health-Related and Non–Health-Related Internet Users

Approximately 59% (265/450, 58.9%) of dementia caregivers were identified as health-related Internet users. Several of the caregiver's characteristics were statistically different between health-related Internet users and non–health-related Internet users. Health-related Internet users were younger ($P$=.01), were more educated ($P$<.001), had a higher level of household income ($P$<.001), and spent fewer hours per week caregiving ($P$=.004). Health-related Internet users were more likely to be a child or grandchild of persons with dementia (78.4%, 207/265) rather than their spouse (3.4%, 9/265) compared to non–health-related Internet users (child or grandchild: 70.0%, 128/185; spouses: 10.9%, 20/185; $P$=.02). However, the sociodemographic characteristics of persons with dementia were not statistically different whether their caregivers were health-related Internet users or not (Table 1).

### Overall Model of a Hierarchical Binary Logistic Regression

The results of the hierarchical binary logistic regression analysis are shown in Table 2. The overall model explained 23.9% of the variance to predict who health-related Internet users were. The group classifications predicted 80.3% of the health-related Internet user group and 55.4% of the non–health-related Internet user group (Tables 2 and 3).

In Block 1, the age and gender of persons with dementia explained 1.3% of the variance (Nagelkerke $R^2$=.013; $\chi^2_2$=3.7; $P$=.15). Both the age and gender of persons with dementia were not significant factors. After adding caregiver's sociodemographic factors into Block 2, the overall model became significant because Block 2 added 16.9% of the explained variance (Nagelkerke $R^2$=.182; $\chi^2_{15}$=54.4; $P$<.001). Caregiver's age (Wald $\chi^2_1$=4.5; $P$=.03; OR 0.980, 95% CI 0.963-0.998) and education levels (Wald $\chi^2_1$=25.8; $P$<.001; OR 3.523, 95% CI 2.168-5.726) were significant factors. Block 3 including ADLs and IADLs did not significantly increase explained variance (Nagelkerke $R^2$=.003; $\chi^2_2$=0.8; $P$=.67). However, overall model and predictive values of caregiver's age and education levels still remained significant ($P$<.05 for all). Block 4 included the number of hours and duration of caregiving years which did not significantly increase explained variance (Nagelkerke $R^2$=.016; $\chi^2_2$=5.1; $P$=.08). However, the overall model including Blocks 1 to 4 still remained significant (Nagelkerke $R^2$=.201; $\chi^2_{19}$=60.3; $P$<.001). There were noticeable changes in individual factors. Caregiver's age became an insignificant factor (Wald $\chi^2_1$=2.5; $P$=.11). After controlling for sociodemographics of persons with dementia and their caregivers, IADLs became a significant factor (Wald $\chi^2_1$=4.0; $P$=.045; OR 1.201, 95% CI 1.004-1.436). After controlling for functional dependency of persons with dementia and all sociodemographics, the number of hours for caregiving was a significant factor (Wald $\chi^2_1$=4.8; $P$=.03; OR 0.519, 95% CI 0.288-0.933). However, caregiver's education levels remained a factor with similar predictive strength (Wald $\chi^2_1$=25.1; $P$<.001; OR 3.536, 95% CI 2.158-5.794).

The final model including Blocks 1 to 5 significantly explained 23.9% of the variance (Nagelkerke $R^2$=.239; $\chi^2_{22}$=73.0; $P$<.001) with Block 5′s significant increase of 2.0% of explained variance ($\chi^2_3$=12.7; $P$=.01). Caregiver's age still remained an insignificant factor (Wald $\chi^2_1$=3.4; $P$=.07). However, caregiver's education levels remained a factor with similar predictive strength (Wald $\chi^2_1$=21.9; $P$<.001; OR 3.348, 95% CI 2.0-5.552). After controlling for sociodemographics of persons with dementia and their caregivers, IADLs became an insignificant factor (Wald $\chi^2_1$=3.2; $P$=.07; OR 1.183, 95% CI 0.985-1.420). After controlling for functional dependency of persons with dementia and all sociodemographic factors, the number of hours for caregiving remained a significant factor, but decreasing in strength of prediction (Wald $\chi^2_1$=6.3; $P$=.01; OR 0.452, 95% CI 0.243-0.840). Newly emerging significant factors were identified. Age of persons with dementia was shown as a significant factor (Wald $\chi^2_1$=4.5; $P$=.03; OR 0.278, 95% CI 0.085-0.906). Caregiver's emotional stress (Wald $\chi^2_1$=4.0; $P$=.046; OR 1.249, 95% CI 1.004-1.555) and caregiver's financial hardship (Wald $\chi^2_1$=6.4; $P$=.01; OR 4.606, 95% CI 1.416-14.978) were significant factors after controlling for sociodemographics, caregiving history, and functional dependency of persons with dementia.

**Table 1.** Sociodemographic and caregiving-related characteristics of persons with dementia and their caregivers.

| Variables | All<br>N=450 | Health-related Internet users<br>(n=265) | Non–health-related Internet users<br>(n=185) | *P* |
|---|---|---|---|---|
| **Description of persons with dementia** [a] | | | | |
| Age (years), mean (SD)[b] | 78.37 (14.1) | 79.38 (12.8) | 76.90 (15.8) | .07 |
| Impairment of ADLs, mean (SD)[c] | 2.21 (2.0) | 2.14 (2.1) | 2.32 (1.9) | .36 |
| Impairment of IADLs, mean (SD)[c] | 4.21 (1.7) | 4.22 (1.7) | 4.18 (1.8) | .80 |
| Gender (female), n (%)[d] | 309 (69.1) | 180 (68.2) | 129 (70.5) | .60 |
| **Description of their caregivers** | | | | |
| Age (years), mean (SD)[c] | 50.30 (15.0) | 48.67 (13.3) | 52.62 (16.8) | .009 |
| Number of hours for caregiving (per week), mean (SD)[b] | 29.96 (46.9) | 26.10 (44.3) | 35.57 (50.1) | .004 |
| Duration of caregiving (years), mean (SD)[b] | 5.22 (8.0) | 4.65 (6.6) | 6.05 (9.6) | .35 |
| Gender (female), n (%)[d] | 277 (62.0) | 156 (58.9) | 121 (66.1) | .12 |
| **Race, n (%)** [d] | | | | .28 |
| Non-Hispanic white | 320 (71.6) | 195 (73.9) | 125 (68.3) | |
| Non-Hispanic African American | 49 (11.0) | 24 (9.1) | 25 (13.7) | |
| Hispanic | 53 (11.9) | 27 (10.2) | 26 (14.2) | |
| Non-Hispanic Asian | 12 (2.7) | 7 (2.7) | 5 (2.7) | |
| Missing data | 16 (2.8) | 12 (4.1) | 4 (1.1) | |
| **Education levels, n (%)** [d] | | | | <.001 |
| High school or less | 240 (53.5) | 109 (40.9) | 131 (71.6) | |
| Some college or higher | 208 (46.5) | 156 (59.1) | 52 (28.4) | |
| **Residence area, n (%)** [d] | | | | .05 |
| Urban | 131 (29.3) | 82 (31.1) | 49 (26.8) | |
| Suburban | 169 (37.8) | 108 (40.9) | 61 (33.3) | |
| Rural | 137 (30.7) | 70 (26.5) | 67 (36.6) | |
| Missing data | 13 (2.2) | 5 (1.5) | 8 (3.3) | |
| **Income (US $), n (%)** [d] | | | | <.001 |
| <$30,000/year | 94 (21.0) | 39 (14.8) | 55 (30.1) | |
| ≥$30,000/year | 314 (70.3) | 201 (76.1) | 113 (61.8) | |
| Missing data | 42 (8.7) | 25 (9.1) | 17 (8.1) | |
| **Relationship to person with dementia, n (%)** [d] | | | | .02 |
| Spouse | 29 (6.5) | 9 (3.4) | 20 (10.9) | |
| Parent | 15 (3.4) | 7 (2.7) | 8 (4.3) | |
| Child or grandchild | 335 (74.9) | 207 (78.4) | 128 (70.0) | |
| Other type of relative | 38 (8.5) | 22 (8.3) | 16 (8.6) | |
| Friend/nonrelative/neighbor | 29 (6.5) | 18 (6.8) | 11 (6.0) | |
| Missing data | 4 (0.2) | 2 (0.4) | 2 (0.2) | |
| **Subjective responses of caregiving stress** | | | | |
| Physical strain, mean (SD)[c] | 2.42 (1.4) | 2.36 (1.3) | 2.51 (1.5) | .27 |
| Emotional stress, mean (SD)[c] | 3.13 (1.3) | 3.20 (1.3) | 3.03 (1.3) | .19 |

| Variables | All<br>N=450 | Health-related Internet users<br>(n=265) | Non–health-related Internet users<br>(n=185) | *P* |
|---|---|---|---|---|
| Financial hardship, mean (SD)[b] | 2.05 (1.3) | 2.08 (1.3) | 2.01 (1.3) | .51 |

[a] ADL=activities of daily living; IADL=instrumental activities of daily living.

[b] Tested by Mann-Whitney *U* tests.

[c] Tested by independent *t* test.

[d] Tested by chi-square test.

**Table 2.** Final model of a hierarchical binary logistic regression analysis to predict health-related Internet users.

| Factors | B (SE) | Wald $\chi^2$ (*df*) | *P* | OR (95% CI) |
|---|---|---|---|---|
| **Block 1: Demographics of persons with dementia** | | | | |
| Constant | 1.461 (1.320) | 1.2 (1) | .27 | 4.309 |
| Age[a] | –1.279 (0.602) | 4.5 (1) | .03 | 0.278 (0.085-0.906) |
| Female gender[b] | 0.102 (0.276) | 0.1 (1) | .71 | 1.108 (0.646-1.901) |
| **Block 2: Sociodemographics of caregivers** | | | | |
| Age | –0.019 (0.010) | 3.7 (1) | .07 | 0.981 (0.962-1.001) |
| Female gender[b] | –0.365 (0.250) | 2.1 (1) | .15 | 0.694 (0.425-1.134) |
| Education levels[c] | 1.208 (0.258) | 21.9 (1) | <.001 | 3.348 (2.019-5.552) |
| Household income[d] | 0.593 (0.318) | 3.5 (1) | .06 | 1.809 (0.969-3.377) |
| Race and ethnicity | | 0.6 (3) | .89 | |
| Resident area | | 1.7 (2) | .42 | |
| Relationship to dementia persons | | 2.8 (4) | .59 | |
| **Block 3: Functional dependency** | | | | |
| ADLs | –0.010 (0.070) | 0.02 (1) | .88 | 0.990 (0.862-1.136) |
| IADLs | 0.168 (0.093) | 3.2 (1) | .07 | 1.183 (0.985-1.420) |
| **Block 4: Caregiving history** | | | | |
| Number of hours for caregiving[a] | –0.795 (0.317) | 6.3 (1) | .01 | 0.452 (0.243-0.840) |
| Duration of caregiving | –0.023 (0.020) | 1.2 (1) | .27 | 0.978 (0.939-1.017) |
| **Block 5: Subjective responses of caregiving stress** | | | | |
| Physical strain | –0.100 (0.117) | 0.7 (1) | .39 | 0.905 (0.719-1.138) |
| Emotional stress | 0.222 (0.112) | 4.0 (1) | .05 | 1.249 (1.004-1.555) |
| Financial hardship[a] | 1.527 (0.602) | 6.4 (1) | .01 | 4.606 (1.416-14.978) |

[a] Transformed using a $\log_{10}$ function.

[b] Reference: male.

[c] Reference: those educated at the level of high school or less.

[d] Reference: those who had household incomes less than US $30,000 per year.

**Table 3.** Odds ratio changes of significant factors.

| Factors | Block 1 | | Block 1 and 2 | | Block 1 to 3 | | Block 1 to 4 | | Block 1 to 5 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | OR | *P* | OR | *P* | OR | *P* | OR | *P* | OR | *P* |
| Constant | 4.84 | .13 | 4.77 | .20 | 4.80 | .21 | 3.07 | .38 | 4.31 | .27 |
| **Block 1: Demographics of persons with dementia** | | | | | | | | | | |
| Age[a] | | | | | | | | | 0.28 | .03 |
| Female gender[b] | | | | | | | | | | |
| **Block 2: Sociodemographics of caregivers** | | | | | | | | | | |
| Age | — | | 0.98 | .17 | 0.98 | .02 | 0.99 | .13 | 0.98 | .07 |
| Female gender[b] | — | | | | | | | | | |
| Education levels[c] | — | | 3.52 | <.001 | 3.54 | <.001 | 3.54 | <.001 | 3.35 | <.001 |
| Household income[d] | — | | | | | | | | | |
| Race and ethnicity | — | | | | | | | | | |
| Resident area | — | | | | | | | | | |
| Relationship with dementia persons | — | | | | | | | | | |
| **Block 3: Functional dependency** | | | | | | | | | | |
| ADLs | — | | — | | | | | | | |
| IADLs | — | | — | | | | 1.20 | .04 | 1.18 | .07 |
| **Block 4: Caregiving history** | | | | | | | | | | |
| Number of hours for caregiving[a] | — | | — | | — | | 0.52 | .03 | 0.45 | .01 |
| Duration of caregiving | — | | — | | — | | | | | |
| **Block 5: Subjective responses of caregiving stress** | | | | | | | | | | |
| Physical strain | — | | — | | — | | — | | | |
| Emotional stress | — | | — | | — | | — | | 1.25 | .04 |
| Financial hardship[a] | — | | — | | — | | — | | 4.61 | .01 |

[a] Transformed using a $\log_{10}$ function.

[b] Reference: male.

[c] Reference: those educated at the level of high school or less.

[d] Reference: those who had household incomes less than US $30,000 per year.

## Discussion

### Principal Results

This study examined the sociodemographic and caregiving characteristics of health-related Internet users among dementia caregivers. In this study, using 2009 NAC and AARP survey data, 59.1% of dementia caregivers were identified as health-related Internet users. Caregiver's age, levels of education and income, hours spent caregiving each week, and relationship to persons with dementia were univariate factors discriminating the health-related Internet use group from non-health-related users. After controlling for confounding effects, age and dependency of IADLs of persons with dementia, caregiver's emotional stress, and caregiver's financial hardship were newly emerging factors of health-related Internet use. Caregiver's sociodemographics and their subjective responses of caregiving stress were the most significant factors to identify health-related Internet users followed by workload assisting in IADLs of persons with dementia [31].

### Comparison With Prior Work

In all, 59% of the prevalence of health-related Internet use is lower than that of the general public (80% in the 2010 Pew Internet & American Life Project) [21], but similar to those found in other types of caregivers (eg, 42%-60% for cancer caregivers) [17]. Compared to findings reported in the recent Pew Internet Health Tracking Survey (2012), our dementia caregiving participants reported a much higher rate (59%) than 39% reported by general caregivers of an adult or child with significant health issues [32]. This finding suggests that using the Internet has become more prevalent and significant behavior seeking for health-related resources among caregivers in the United States. This intensity seems to result from (1) the huge growth in availability and the widespread adoption of the Internet and relevant technologies [33], (2) health care consumers' strong motivation, (3) their positive perceptions regarding Internet-based resources [34], and (4) the promising benefits of Internet-based approaches (ie, convenience and confidentiality) [34]. Thus, the Internet has been acknowledged as a promising modality for implementing interventions or distributing caregiving resources.

Caregiver's sociodemographic characteristics are strong factors in determining their behavior regarding health-related Internet use, including age, education levels, income, and their relationship to persons with dementia. Similar to previous study findings, the predictive values of age and education levels were confirmed in this study. Younger and more highly educated caregivers reported they used Internet-based resources for better health and caregiving purposes similar to the same findings in the general population or of cancer caregivers [17,33-36]. Interestingly, those who had a minimum of college-level education were 3.35 times more likely to be health-related Internet users than those who were educated at the level of high school or less. Age and educational attainment were the most significant factors for eHealth literacy [37,38]. In addition, younger generations were generally considered to be more technology-friendly and more prepared to use Internet-based resources compared to their older counterparts. Additionally, a higher education level may be associated with either a higher level of knowledge of health-related resources or better computer skills [15,38].

Higher levels of both household income and self-reported financial hardship were associated with a greater likelihood of being health-related Internet users. Initially, those 2 findings may appear to conflict with one other because a basic assumption is that a person with higher income would experience a lower level of self-reported financial hardship. However, this inconsistency has been reported in previous research; as with the study findings, individuals with higher levels of income use the Internet more [33,35]. Individuals with higher household incomes are more likely to own computers and handheld mobile devices and spend more time using the Internet in their daily life than their lower-income counterparts [33,37]. However, another study has described the opposite association between income and health-related Internet use. A study of national surveys reported that those with lower incomes were also more likely to participate in online support groups than those with higher incomes [36]. Interestingly, this study found the unique impact of the subjective response of financial hardship after controlling for income levels. One possible explanation is that individuals with higher incomes may have diverse means to reduce their stress levels. However, those who are experiencing financial hardship have limited use of traditional face-to-face resources for stress reduction and they try to find alternatives through the Internet [33].

Spousal caregivers were significantly less likely to use available services, which is consistent with previous reports [14]. This reluctance by spousal caregivers seems to originate from emotional barriers: spouses tend to perform the caregiving role without external help because using resources may be considered as betraying the spousal relationship [14]. However, the significant impact of their relationship seems to result from the confounding effect of their age. After controlling for their age, the predictive value is no longer significant in this study sample. Additionally, children of persons with dementia would naturally be considerably younger than their parent and, thus, likely to be more technology-prepared [39].

One of the salient findings of this study was that the higher the emotional stress experienced by dementia caregivers, the more health-related Internet use they reported. When they reported very much emotional stress (score 5) on a 5-point Likert-type scale, they were 3.05 times more likely to be health-related Internet users than those who reported no emotional stress (score 1). This tendency was found in a previous study that applied qualitative methods to the online postings of dementia caregivers. The emotional concern and seeking psychosocial support represents the second-commonest theme for dementia caregivers after behavior of seeking information [20]. Caregiver stress has been shown as a need variable that facilitates their use of resources [40], especially in dementia caregivers [6,15,40]. Dementia caregivers with higher levels of emotional and psychological stress are more likely to use the traditional face-to-face resources of health and human services [15]. The stress-appraisal theory [41] and the stress process model [13] explain that there is a positive relationship between recognized stress levels and efforts to alleviate stress. In addition, a systemic review of networked technologies in dementia caregiver stated a potential impact of ITC intervention on caregiver stress [42]. Thus, health-related Internet use may be considered a coping strategy for caregivers to relieve their subjective stress or burden [41] and a mediator to modulate the impact of caregiving stress on negative outcomes [13].

The predictive value of functional dependency of IADLs appears after controlling for sociodemographic impact. When caregivers are taking care of totally dependent persons with dementia (IADLs score=6), they were 2.99 times less likely to be health-related Internet users than those taking care of totally independent persons with dementia (IALDs score=0). However, when adding the number of hours for caregiving, the predictive value of IADLs dependency became insignificant. The fewer hours caregivers spent caregiving, the more likely they were to be health-related Internet users. For those caring for cancer survivors, having fewer direct caring responsibilities increases the likelihood of health-related Internet use [17]. When caregivers spent longer times providing direct caregiving tasks, they did not have the time or energy to search for information and seek support via the Internet. Respite care may be especially beneficial to caregivers because it may free up the time they need to access the Internet for health-related purposes. Thus, this finding may suggest more appropriate ways to apply interventions designed for providing physical assistance and, thereby, reducing the workload of dementia caregivers.

### Study Limitations and Directions for Future Research

This study had several limitations: (1) limited inference of causality due to the cross-sectional design, (2) limited generalizability due to the use of a convenience sampling method, (3) possible responder bias in the self-reported data, and (4) difficulties in controlling data quality as a secondary data analysis.

Conceptually, dichotomous grouping (health-related Internet users vs non–health-related Internet users) has limitations in explaining their complex behavior of seeking resources for caregiving. Using current data, the multiple grouping depending on intensity of use (eg, high vs moderate vs nonuser groups) and multinomial logistic regression may show a more detailed description in this sample. Moreover, data collected in 2009 limitedly reflect current trends in the scientific community or daily practice because Internet research is 1 of the most rapidly changing in the field.

Future up-to-date studies would greatly benefit from the use of a longitudinal design that utilizes other multiple measurements of health-related Internet use or caregiver stress. Because the variables tested in this study are predisposing and need variables to facilitate health care service use [14], the inclusion of enabling variables related to health service utilization should provide a contextually better in-depth understanding to shed new light on the complex picture of health-related Internet use in this population. To define the medical condition of dementia using International Classification of Diseases codes will be more accurate than caregiver-reported condition.

## Conclusions

The Internet has become a significant resource for dementia caregivers for health-related purposes. This research adds to our knowledge of the prevalence and factors of health-related Internet use by dementia caregivers. Subjective responses of caregiving stress are a need factor leading to increased Internet use for health and caregiving purposes. Significant demographic factors provide helpful information to identify those who are less likely to use Internet-based resources. The lowest utilization is detected in those who were older, a spouse, less educated, with lower incomes, and devoting longer times to caregiving. Thus, this study helps us identify underserved groups regarding virtual health care resources. Clinical researchers should consider our findings to develop tailored interventions and effective care delivery approaches targeting the virtually underserved caregiving population. Additionally, the study findings may assist policy makers seeking to distribute information, resources, and services via the Internet to help dementia caregivers and their care recipients with dementia.

## Acknowledgments

The author appreciates the generosity of the National Alliance for Caregiving and the American Association of Retired Persons in allowing us to use their data, published in 2010. The data from the NAC and AARP’s National Caregiver survey are available in electronic form to nonprofit research organizations and individuals and deidentifiable for public use at no charge. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Conflicts of Interest

None declared.

## References

1. Centers for Disease Control and Prevention. 2013. Dementia/Alzheimer's Disease URL: http://www.cdc.gov/mentalhealth/basics/mental-illness/dementia.htm [accessed 2014-07-09] [WebCite Cache ID 6QwnS0XiJ]
2. Schulz R, Martire LM. Family caregiving of persons with dementia: prevalence, health effects, and support strategies. Am J Geriatr Psychiatry 2004;12(3):240-249. [Medline: 15126224]
3. Alzheimer's Association. 2014 Alzheimer’s Disease Facts and Figures. 2014. URL: http://www.alz.org/downloads/facts_figures_2014.pdf [accessed 2014-07-09] [WebCite Cache ID 6QwnvR5uJ]
4. Papastavrou E, Kalokerinou A, Papacostas SS, Tsangari H, Sourtzi P. Caring for a relative with dementia: family caregiver burden. J Adv Nurs 2007 Jun;58(5):446-457. [doi: 10.1111/j.1365-2648.2007.04250.x] [Medline: 17442030]
5. Serrano-Aguilar PG, Lopez-Bastida J, Yanes-Lopez V. Impact on health-related quality of life and perceived burden of informal caregivers of individuals with Alzheimer's disease. Neuroepidemiology 2006;27(3):136-142. [doi: 10.1159/000095760] [Medline: 16974108]
6. Lu YF, Wykle M. Relationships between caregiver stress and self-care behaviors in response to symptoms. Clin Nurs Res 2007 Feb;16(1):29-43. [doi: 10.1177/1054773806295238] [Medline: 17204806]
7. Mannion E. Alzheimer's disease: the psychological and physical effects of the caregiver's role. Part 2. Nurs Older People 2008 May;20(4):33-38. [doi: 10.7748/nop2008.05.20.4.33.c8221] [Medline: 18549115]
8. Christakis NA, Allison PD. Mortality after the hospitalization of a spouse. N Engl J Med 2006 Feb 16;354(7):719-730. [doi: 10.1056/NEJMsa050196] [Medline: 16481639]
9. Mehta A, Cohen SR, Chan LS. Palliative care: a need for a family systems approach. Palliat Support Care 2009 Jun;7(2):235-243. [doi: 10.1017/S1478951509000303] [Medline: 19538807]
10. Cooper C, Blanchard M, Selwood A, Walker Z, Livingston G. Family carers' distress and abusive behaviour: longitudinal study. Br J Psychiatry 2010 Jun;196(6):480-485 [FREE Full text] [doi: 10.1192/bjp.bp.109.071811] [Medline: 20513860]
11. Yaffe K, Fox P, Newcomer R, Sands L, Lindquist K, Dane K, et al. Patient and caregiver characteristics and nursing home placement in patients with dementia. JAMA 2002 Apr 24;287(16):2090-2097. [Medline: 11966383]

12. McClendon MJ, Smyth KA, Neundorfer MM. Survival of persons with Alzheimer's disease: caregiver coping matters. Gerontologist 2004 Aug;44(4):508-519. [Medline: 15331808]
13. Pearlin LI, Mullan JT, Semple SJ, Skaff MM. Caregiving and the stress process: an overview of concepts and their measures. Gerontologist 1990 Oct;30(5):583-594. [Medline: 2276631]
14. Robinson KM, Buckwalter KC, Reed D. Predictors of use of services among dementia caregivers. West J Nurs Res 2005 Mar;27(2):126-40; discussion 141. [doi: 10.1177/0193945904272453] [Medline: 15695566]
15. Toseland RW, McCallion P, Gerber T, Banks S. Predictors of health and human services use by persons with dementia and their family caregivers. Soc Sci Med 2002 Oct;55(7):1255-1266. [Medline: 12365535]
16. National Alliance for Caregiving, American Association of Retired Persons (NAC & AARP). Caregiving in the US 2009 Appendix A: Full questionnaire URL: http://www.caregiving.org/pdf/research/2009Caregiving_Appendix_A_Full_Questionnaire.pdf [accessed 2012-01-26] [WebCite Cache ID 64zHNc61Z]
17. Kinnane NA, Milne DJ. The role of the Internet in supporting and informing carers of people with cancer: a literature review. Support Care Cancer 2010 Sep;18(9):1123-1136. [doi: 10.1007/s00520-010-0863-4] [Medline: 20336326]
18. Schulz R, Lustig A, Handler S, Martire L. Technology-Based Caregiver Intervention Research: Current Status and Future Directions. Gerontechnology 2002 Jan 01;2(1). [doi: 10.4017/gt.2002.02.01.003.00]
19. Lewis ML, Hobday JV, Hepburn KW. Internet-based program for dementia caregivers. Am J Alzheimers Dis Other Demen 2010 Dec;25(8):674-679. [doi: 10.1177/1533317510385812] [Medline: 21131674]
20. White MH, Dorman SM. Online support for caregivers. Analysis of an Internet Alzheimer mailgroup. Comput Nurs 2000;18(4):168-76; quiz 177. [Medline: 10939186]
21. Fox S, Brenner J. Pew Research Internet Project. 2012 Jul 12. Family caregivers online URL: http://www.pewinternet.org/Reports/2012/Caregivers-online.aspx [accessed 2013-11-22] [WebCite Cache ID 6LJtK1k7N]
22. National Caregiving Alliance and United Healthcare. 2011 Jan. e-Connected family caregivers: Bringing caregiving into the 21st century URL: http://www.caregiving.org/data/FINAL_eConnected_Family_Caregiver_Study_Jan%202011.pdf [accessed 2012-11-26] [WebCite Cache ID 6CTLDtD51]
23. Lauriks S, Reinersmann A, Van der Roest HG, Meiland FJ, Davies RJ, Moelaert F, et al. Review of ICT-based services for identified unmet needs in people with dementia. Ageing Res Rev 2007 Oct;6(3):223-246. [doi: 10.1016/j.arr.2007.07.002] [Medline: 17869590]
24. National Alliance for Caregiving, American Association of Retired Persons (NAC & AARP). Appendix B: Detailed methodology URL: http://www.caregiving.org/data/2009Caregiving_Appendix_B_Detailed_Methodology.pdf [accessed 2012-01-26] [WebCite Cache ID 64zGa13u2]
25. National Alliance for Caregiving, American Association of Retired Persons. Caregiving in the US: Executive summary URL: http://www.caregiving.org/pdf/research/CaregivingUSAllAgesExecSum.pdf [accessed 2012-01-25] [WebCite Cache ID 64zHKfezm]
26. Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standarized meausre of biological and psychosocial function. JAMA 1963;185:914-919. [Medline: 14044222]
27. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-186. [Medline: 5349366]
28. Meyers LS, Gamst G, Guarino AJ. Applied Multivariate Research: Design and Interpretation. Thousand Oaks, CA: Sage Publications; 2006.
29. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]
30. Wilson CR, Voorhis V, Morgan BL. Understanding power and rules of thumb for determining sample sizes. Tutorials in Quantitative Methods for Psychology 2007;3920:43-50.
31. Zulman DM, Piette JD, Jenchura EC, Asch SM, Rosland AM. 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]
32. Fox S, Duggan M, Purcell K. Pew Research Internet Project. 2013 Jun 20. Family caregivers are wired for health URL: http://www.pewinternet.org/2013/06/20/family-caregivers-are-wired-for-health/ [accessed 2014-07-13] [WebCite Cache ID 6R3097Y8r]
33. National Telecommunications and Information Administration. Exploring the digital nation computer and Internet use at home. 2011 Nov. URL: http://www.ntia.doc.gov/files/ntia/publications/exploring_the_digital_nation_computer_and_internet_use_at_home_11092011.pdf [accessed 2012-01-26] [WebCite Cache ID 64zIATlo8]
34. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
35. Cresci MK, Yarandi HN, Morrell RW. The digital divide and urban older adults. Comput Inform Nurs 2010 Apr;28(2):88-94. [doi: 10.1097/NCN.0b013e3181cd8184] [Medline: 20182159]
36. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]

37. Choi NG, Dinitto DM. The digital divide among low-income homebound older adults: Internet use patterns, eHealth literacy, and attitudes toward computer/Internet use. J Med Internet Res 2013;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
38. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
39. Choi N. Relationship between health service use and health information technology use among older adults: Analysis of the US national health interview survey. J Med Internet Res 2011;Apr [FREE Full text] [doi: 10.2196/jmir.1753]
40. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
41. Lazarus RS, Folkman S. Stress, Appraisal, and Coping. New York: Springer; 1984.
42. Powell J, Chiu T, Eysenbach G. A systematic review of networked technologies supporting carers of people with dementia. J Telemed Telecare 2008;14(3):154-156. [doi: 10.1258/jtt.2008.003018] [Medline: 18430288]

## Abbreviations

**AARP:** American Association of Retired Persons
**ADLs:** activities of daily living
**IADLs:** instrumental activities of daily living
**ICT:** information and communication technology
**NAC:** National Alliance for Caregiving

*Edited by G Eysenbach; submitted 22.11.13; peer-reviewed by B Hattink, AM Pot; comments to author 28.04.14; revised version received 13.07.14; accepted 22.07.14; published 23.02.15.*

*Please cite as:*
*Kim H*
*Understanding Internet Use Among Dementia Caregivers: Results of Secondary Data Analysis Using the US Caregiver Survey Data*
*Interact J Med Res 2015;4(1):e1*
*URL: http://www.i-jmr.org/2015/1/e1/*
*doi:10.2196/ijmr.3127*
*PMID:25707033*

Original Paper

# Real-Time Web-Based Assessment of Total Population Risk of Future Emergency Department Utilization: Statewide Prospective Active Case Finding Study

Zhongkai Hu[1,2*], BS; Bo Jin[3*], MS; Andrew Y Shin[4*], MD; Chunqing Zhu[3], MS; Yifan Zhao[3], MS; Shiying Hao[2], PhD; Le Zheng[2,5], MS; Changlin Fu[3], MS; Qiaojun Wen[3], PhD; Jun Ji[3,6], PhD; Zhen Li[2,7], MS; Yong Wang[8,9], PhD; Xiaolin Zheng[1,2], PhD; Dorothy Dai[3], BS; Devore S Culver[10], MBA; Shaun T Alfreds[10], MBA; Todd Rogow[10], MPA; Frank Stearns[3], MHA; Karl G Sylvester[2], MD; Eric Widen[3], MHA; Xuefeng B Ling[2], PhD

[1]College of Computer Science and Technology, Zhejiang University, Hangzhou, China
[2]Department of Surgery, Stanford University, Stanford, CA, United States
[3]HBI Solutions Inc, Palo Alto, CA, United States
[4]Department of Pediatrics, Stanford University, Stanford, CA, United States
[5]Department of Electrical Engineering, Tsinghua University, Beijing, China
[6]College of Information Engineering, Qingdao University, Qingdao, China
[7]Institute of Microanalytical System, Zhejiang University, Hangzhou, China
[8]Department of Statistics, Stanford University, Stanford, CA, United States
[9]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China
[10]HealthInfoNet, Portland, ME, United States
[*]these authors contributed equally

**Corresponding Author:**
Xuefeng B Ling, PhD
Department of Surgery
Stanford University
300 Pasteur Dr
Stanford, CA,
United States
Phone: 1 650 427 9198
Fax: 1 650 723 1154
Email: bxling@stanford.edu

## Abstract

**Background:** An easily accessible real-time Web-based utility to assess patient risks of future emergency department (ED) visits can help the health care provider guide the allocation of resources to better manage higher-risk patient populations and thereby reduce unnecessary use of EDs.

**Objective:** Our main objective was to develop a Health Information Exchange-based, next 6-month ED risk surveillance system in the state of Maine.

**Methods:** Data on electronic medical record (EMR) encounters integrated by HealthInfoNet (HIN), Maine's Health Information Exchange, were used to develop the Web-based surveillance system for a population ED future 6-month risk prediction. To model, a retrospective cohort of 829,641 patients with comprehensive clinical histories from January 1 to December 31, 2012 was used for training and then tested with a prospective cohort of 875,979 patients from July 1, 2012, to June 30, 2013.

**Results:** The multivariate statistical analysis identified 101 variables predictive of future defined 6-month risk of ED visit: 4 age groups, history of 8 different encounter types, history of 17 primary and 8 secondary diagnoses, 8 specific chronic diseases, 28 laboratory test results, history of 3 radiographic tests, and history of 25 outpatient prescription medications. The c-statistics for the retrospective and prospective cohorts were 0.739 and 0.732 respectively. Integration of our method into the HIN secure statewide data system in real time prospectively validated its performance. Cluster analysis in both the retrospective and prospective analyses revealed discrete subpopulations of high-risk patients, grouped around multiple "anchoring" demographics and chronic conditions. With the Web-based population risk-monitoring enterprise dashboards, the effectiveness of the active case finding algorithm has been validated by clinicians and caregivers in Maine.

**Conclusions:** The active case finding model and associated real-time Web-based app were designed to track the evolving nature of total population risk, in a longitudinal manner, for ED visits across all payers, all diseases, and all age groups. Therefore, providers can implement targeted care management strategies to the patient subgroups with similar patterns of clinical histories, driving the delivery of more efficient and effective health care interventions. To the best of our knowledge, this prospectively validated EMR-based, Web-based tool is the first one to allow real-time total population risk assessment for statewide ED visits.

***(Interact J Med Res 2015;4(1):e2)*** doi:10.2196/ijmr.4022

**KEYWORDS**

ED; machine learning; HIE; EMR; modeling

## Introduction

The use of emergency department (ED) services in the United States is growing at an alarming rate [1-3]. Between 2001 and 2008, the annual number of ED visits in the United States grew at roughly twice the rate of population growth [4]. Recent experience from Oregon's Health Insurance Experiment suggests that increasing patient access to Medicaid without an accompanying strategy to manage the overall insured population may result in a substantial surge in ED utilization [5], including visits for conditions that may be most readily treatable in primary care settings. Presuming a large proportion of ED visits are preventable, attention has turned toward strategies to treat patients in less expensive outpatient care settings, and payers are beginning to deny payment for non-urgent ED visits [6].

Improving appropriate use of emergency services is an important strategy for improving health outcomes and controlling health care expenditures [7]. With the increased adoption of electronic medical record (EMR) systems and the development of health information exchanges (HIE) in the United States, health care organizations have better and more comprehensive access to patients' comprehensive medical histories. Greater use of advanced analytic computing methods on EMR datasets has led to the development of several active case finding algorithms to assess patient risk. Early efforts included risk prediction models for hospital readmission [8] and repeated ED visits for patients with distinct patterns [9-11]. Most risk development studies focused on patients within specific payer groups, for example, Medicare, within specific age, and/or within specific disease groups [12,13].

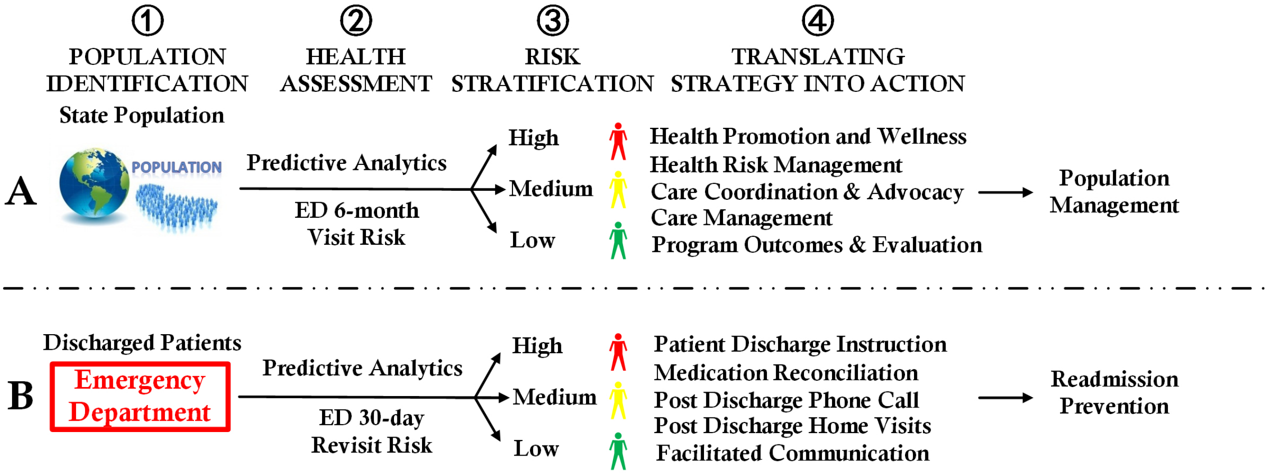
We previously developed predictive analytics of patient risk of a 30-day return to the emergency department [14]. The 30-day ED revisit risk is intended for hospital emergency room and quality management staff to immediately plan for post-discharge care while the patient is in the emergency room, or shortly thereafter. This particular risk is triggered by the event of an emergency room visit, and therefore is a very small subset of the whole population, that is, only those patients with at least one emergency room visit are covered. Second, emergency room revisit rates are a quality measure used to assess hospital performance.

In this paper, we describe our findings for the ED visit risk modeling for the statewide population at large, whether or not they have had a previous emergency room visit. This is the first effort to model total population ED risk across all payers, all diseases, and all age groups. Our efforts include the statistical learnings from all Maine HIE patient data contained in the statewide HIE of longitudinal patterns to identify risk factors that strongly influence the probability of a future 6-month ED visit.

Although the two metrics (ie, risks of the 30-day ED revisit [14] and the future 6-month ED visit), have similarities in regard to ED visit risk, these are two distinct risks for two distinct purposes (Figure 1). We studied both to understand differences and similarities between them. The population 6-month ED visit risk is intended for the care team responsible for population health management in accountable care organizations (ACOs) and providers with capitated risk contracts.

We hypothesized that real-time assessment of population ED risk to track and trend risk over time can allow health managers to continuously assess and intervene on both high-risk and rising-risk patients. To empower the visualization and exploration of the total population risks of over one million patients in the state of Maine, Web-based apps were designed, aiming to connect in real-time, aggregate, and centrally integrate data, and to compute future 6-month ED risks for population health management.

**Figure 1.** Integrating predictive analytics into workflows of proactive population health management and hospital quality improvement; emergency department (ED) visit risk determination and proactive interventions guided by ED visit risk or ED readmission risk measures.



Hao S, Jin B, Shin AY, Zhao Y, Zhu C, et al. (2014) PLOS ONE 9(11): e112944. doi:10.1371

## Methods

### Ethics Statements

This work was done under a business/product development arrangement between HealthInfoNet (HIN) and HBI Solutions, Inc., and the data use was governed by a business agreement between HIN and HBI. No patient health information was released for the purpose of research and no patient consent was required. We completed the system development that was the foundation for our agreement and then reported on the findings resulting from applying this model to the real-time Web-based services that HIN is now deploying in the field. Because this study analyzed de-identified data to develop the ED risk model, the Stanford University Institutional Board considered it exempt (October 16, 2014).

### Population

The objective was to study total population risk for ED visits across all payers, all diseases, and all age groups. Patients visiting any HIN-connected facility from January 1, 2012 through December 31, 2013 were eligible for study. Patients who died, as identified through an encounter disposition code, were excluded during the study time frame of 2012 and 2013. ED visits transferred from another ED were excluded as these were treated as one ED visit, and not multiple.

### Data Acquisition and Marshalling

We constructed an enterprise data warehouse consisting of all of Maine’s HIE aggregated patient histories. Incorporated data elements from EMR encounters included patient demographic information, laboratory tests and results, radiographic procedures, medication prescriptions, diagnosis, and procedures, which were coded according to the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM). Census data from the US Department of Commerce Census Bureau were integrated into our data warehouse. Therefore, in addition to the HIN features, we categorized patients by socioeconomic status using residence zip codes as an approximation to the average household mean and median family income and average degree of educational attainment. Maine HIE patient clinical histories were organized as hospital episode level relational database tables. We processed the database at patient level based on medical record number for population analysis within 36 facilities in Maine. A pivot table was developed from our enterprise data warehouse, which aggregated and integrated normalized clinical features (n=33,403) of different data categories, for example, primary diagnosis/procedure, secondary diagnosis/procedure, laboratory test result, radiology result, and outpatient prescription, from different relational EMR databases. For qualitative and categorical parameters, dummy variables were created serving as numerical representations of the categories of nominal or ordinal variables. To efficiently eliminate the least representative features, we exploited the data variance as the simplest criterion [15], which essentially projected the data points along the dimensions of maximum variances. One potential limitation was that variance alone does not account for parameters that had a small dynamic range. However, as an initial filter, this method effectively eliminated “low information content” features to deliver a manageable feature set, allowing the subsequent machine learning step to identify discriminant features. As a result, a set of patient clinical historical features in the prior 12 months was compiled (Multimedia Appendix 1). One of the key features was whether the patient had a chronic medical condition. This feature was defined using the Agency for Healthcare Research and Quality Chronic Condition Indicator [16], which provides an effective way to categorize ICD-9-CM diagnosis codes into one of two categories: chronic and non-chronic.

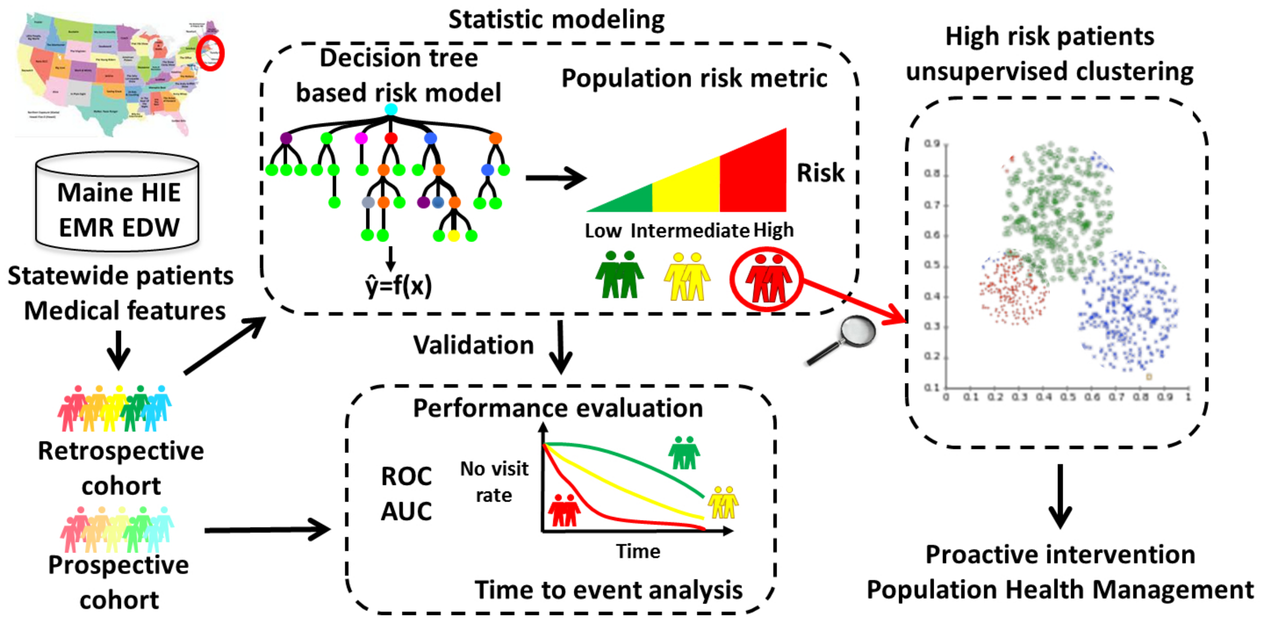
### Outcome Time Frame for Risk Analysis

A “time-to-event” curve of ED visits (Multimedia Appendix 2) was developed to determine whether 6-month ED visit assessment was clinically reasonable. More than 80% of patients with more than one ED visit history would seek ED services within the future 6-month time frame. Therefore, future 6-months was a clinically appropriate cutoff. This was in line with clinical and field caregiver judgments.

### Data Mining Overview: Retrospective and Prospective Analyses

The basic principle of our model was using information of 1 patient in the prior 1 year to predict if this patient would have any ED visit in the next 6 months. The statistical learning to forecast future 6-month ED visit risk consisted of two phases: retrospective modeling and prospective validation (Figure 2). A retrospective cohort of 829,641 patients (Multimedia Appendix 3) who had historical encounter records from January 1 to December 31, 2012, was assembled for the development of the ED risk model to predict if those patients would have ED visits in the next 6 months (January 1 to June 30, 2013). This model was later validated with a prospective cohort of 875,979 patients (Multimedia Appendix 3) who had historical encounter records from July 1, 2012, to June 30, 2013 to predict if these patients would have ED visits in the next 6 months (July 1 to December 31, 2013). Both cohorts of patients had comprehensive clinical histories allowing the determination of future 6-month ED visit risk. Patients in the retrospective and prospective cohorts were similar in age, gender, income, and education, as well as incidence of future 6-month ED visits (retrospective: 11.48%, 95,241/829,641; prospective: 11.37%, 99,558/875,979) (see Multimedia Appendix 4).

**Figure 2.** Study design to develop the active case finding algorithm to predict future 6-month emergency department visit risks.



### Retrospective Analysis Summary

The goal of this study was to develop an active case finding algorithm with a statewide future 6-month ED visit risk measure. The measure comprised a single summary score, derived from the results of a "forest" of the most discriminative decision trees upon 1 year of the encounter history. The measure calculated each subject's probability of a future 6-month ED visit. The retrospective modeling phase consisted of three steps: (1) training, (2) calibrating, and (3) blind testing. We applied a selective cohort division process while trying to result in a random cohort. The samples in the retrospective cohort were divided into six subgroups based on histories of chronic diseases, historical ED visits, and current primary diagnoses (Multimedia Appendix 3). Then, in each subgroup, the case (future 6-month ED visit counts > 0) and control (future 6-month ED visit count=0) samples were randomly partitioned into three cohorts (Cohort I: training cohort, Cohort II: calibrating cohort, Cohort III: blind testing cohort), with the consideration that the past 12-month ED histories of encounters and principle clinical features (chronic diseases and current primary diagnoses) achieved a balance between the cohorts. Therefore, it was hard to achieve a complete balance such that total samples in training, calibrating, and blind testing cohorts had the exact same number. Within each subgroup sharing balanced numbers of chronic histories, ED visits, and current primary diagnoses, the patient numbers in training, calibrating, and blind testing cohorts were close.

### Decision Tree-Based Modeling

A "survival forest" of forecasting decision trees was developed using the prior year clinical history and was ranked according to the corresponding posterior probability. To introduce the prior knowledge, we grouped the clinical features into two groups: empirical features found by exploratory data analysis and the learned features discovered during the model training. Our exploratory analysis (Multimedia Appendix 5) of the retrospective cohort showed that the percentage of patients with future 6-month ED visits increased as a function of either historic ED visit counts or the presence of chronic disease diagnoses; therefore, these two features were strongly associated with patients' risk for future 6-month ED visits. Using empirical features of whether patients had historic ED visit or a chronic disease diagnosis in the prior year, we built a decision tree. This deterministic tree partitioned the Cohort I samples into four subgroups. Within each subgroup, learned features were discovered through the feature selection process to develop the correspondent learning model for the targeted subgroup. Survival tree analysis was applied to learning model process to predict ED visit day after predicted time. Technical details of the model training process [14,17,18] are presented in Multimedia Appendix 6.

### Risk Scoring Metric Development

Cohort II was used to calibrate the predictive scoring threshold to create a risk measure for each individual sample. Applying the model developed with Cohort I to each sample in Cohort II, the derived predictive scores were ranked. After this, we applied a mathematic function mapping predictive values (PPVs).

Our active case finding algorithm was set to segregate the population into subgroups with different levels of future 6-month ED risks. The risk measure was defined as an index between 0 and 100 so that the people with measures larger than or equal to a risk index $L$ had a probability of $L$% to have an ED visit in the next 6 months. Here, the mapped PPV was defined as the individual's risk measure for the future 6-month ED visit.

We obtained two thresholds, $T_h$,$T_m$, from this mapping. The intent of the model was to stratify the patients from low to high risk allowing the clinicians to target different risk levels for personalized intervention. Field care providers can target different risk groups with different threshold settings as a continuous variable for active case finding. Two thresholds of 0.3 and 0.7 were chosen and applied to the ranked outputs of the model to divide the population into low (score<30%), medium (score≥30% and score<70%), and high (score≥70%) risk groups [7].

### Identification of the Discriminant Features

In our implementation, the objective was to select the least number of representative features predictive of future 6-month ED risk and to achieve optimal case finding sensitivity while maintaining the targeted PPV (>70%) based on selected features (Multimedia Appendix 7, left panel). The active case finding algorithm identified 101 variables (Multimedia Appendix 7, right panel) predictive of future 6-month ED risk, which fell into the following general categories: age groups (n=4), history of different encounter types (n=8), history of primary (n=17) and secondary (n=8) diagnosis, specific chronic diseases (n=8), laboratory test results (n=28), history of radiographic tests (n=3), and history of outpatient prescription medications (n=25). The predictive power of the selected features was examined by shrunken difference [19] (Retrospective: Multimedia Appendix 7, right panel; Prospective: Multimedia Appendix 7), which was the scaled distance between the mean values of each feature variable in a specific risk class (low, medium, or high) and across all cohort samples. Shrunken differences among the low-, medium-, and high-risk outcomes differed more than the case (with future ED) and control (without future ED) outcomes, demonstrating the effectiveness of these features in the risk stratification.

### Blind Testing

Cohort III was an independent naive sample set, which was compiled to blind test the active case finding method's performance. The aim of this step was to critically assess the utility of the risk measure before statewide prospective validation in Maine. Again the model developed with Cohort I was applied to every sample in Cohort III to derive and rank the predictive scores and calculated the receiver operating characteristic (ROC) area under the curve (AUC) score.

### Prospective Validation

The clinical application of the future 6-month risk measure was deployed for prospective validation on the HIE data in Maine. The cohort of 875,979 patients from July 1, 2012 to June 30, 2013 was prospectively profiled to calculate future 6-month ED visit risk measures using the clinical applications deployed at HIN. The ROC [20] and time-to-event analyses were performed to gauge the model performance (Multimedia Appendix 8) and effectiveness of the risk stratification.

### Unsupervised Clustering: Subgroup Analysis of High-Risk Patients

We used principle component analysis [21] to reduce high dimensional EMR features and identify clinically relevant groups of patients of high risk for 6-month ED visit with similar patterns of demographics, primary diagnosis and procedure, and chronic disease conditions. The features for high-risk patients were projected to a lower dimensional subspace with largest variances. The $K$-means algorithm was applied to find potential patient patterns for future 6-month ED visit [22]. We used $K$=6 to generate the final six clusters. The technical details are described in Multimedia Appendix 9. Clustering patterns between retrospective and prospective cohorts were compared to further validate our high-risk case finding algorithm. As part of the health care management platform, our predictive model was integrated onto a Web-based dashboard to provide a real-time visualization of the population profile with ED 6-month visits.

### Population Explorer Service: Statewide Real-Time Surveillance of Population ED Risks

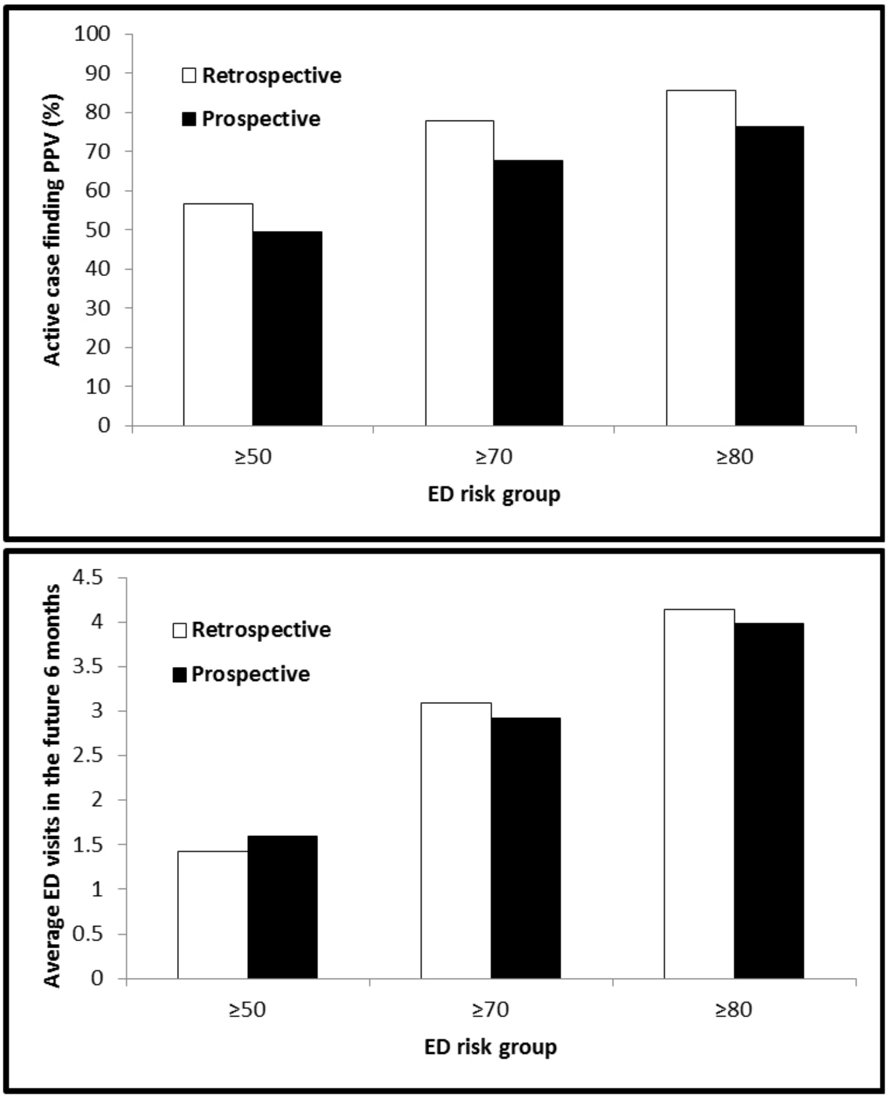
The active case finding model and associated real-time Web-based app were designed to track the evolving nature of total population risk, in a longitudinal manner, for ED visits across all payers, all diseases, and all age groups. Patient historical datasets are linked and stored in a patient-level database in our system. ED predictive algorithm is applied to the individual's ED discriminating feature data to risk-stratify the patients with our prospectively validated model. Individual data are then aggregated for population exploration of ED risks, which are stored in the population-level database. Our dashboard allows the visualization of the population ED risks at high geographical resolution for a defined population, for example, the population of Maine.

## Results

### Data Mining Overview: Retrospective and Prospective Analyses

The active case finding algorithm produced a risk score (from 0 to 100) for each patient at the time of risk assessment of the future ED visit. In general, our algorithm achieved high performance that ROC AUCs of the risk score for a determination of risk of patient future 6-month ED utilization were 0.739 and 0.696 in retrospective blind testing and prospective validating respectively (Multimedia Appendix 8). Specifically, at a risk score threshold of 50, the active case finding algorithm identified 56.55% (9459/16,727) of retrospective and 49.35% (10,810/21,904) of prospective patients who had an ED visit in the next 6 months; 43.45% (7268/16,727) of retrospective or 50.65% (11,094/21,904) of prospective patients were identified incorrectly (who did not have an ED visit) (Figure 3). At risk score threshold levels of 70 and 80, the rate of incorrectly "flagged" patients dropped to 22.20% (839/3780) and 14.48% (286/1975) in retrospective, and 32.31% (1764/5460) and 23.69% (626/2642) in prospective analysis respectively, but the algorithm found a lower percentage of patients. The ROC analyses showed that there was a 0.739 (retrospective) or 0.732 (prospective) probability that a randomly selected patient with a future 6-month ED visit would receive a higher-risk score than a randomly selected patient who did not have a future 6-month ED visit.

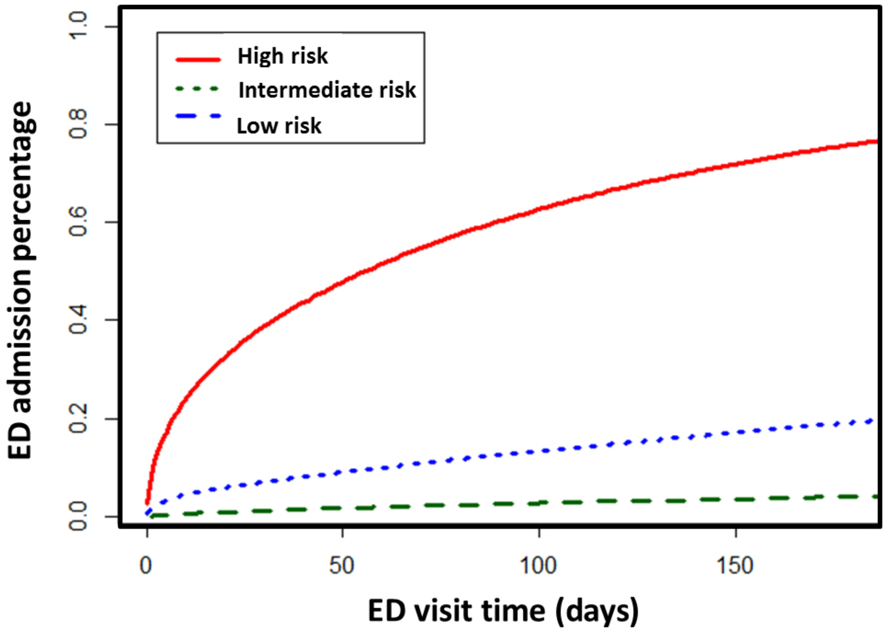
**Figure 3.** Active case finding algorithm effectively identified different risk group patients for future 6-month emergency (ED) utilization (upper panel shows X axis: different ED risk groups; Y axis: active case finding positive predictive value (PPV); and lower panel summarizes average ED uses at different ED risks in the future 6 months in both retrospective and prospective analyses).



## Prospective Validation

In developing the algorithm, we aimed to help potential care providers assess the "opportunity case" (high-cost, high degree of utilization of services, multiple chronic conditions) for various risk score thresholds and for different assumptions about the impact of the intervention. The active case finding algorithm was capable of stratifying patients across a wide range of risks (Figure 3, upper panel) and demonstrated that patients in higher-risk categories visited the ED earlier (prospective time-to-event analysis: $P<.001$) both on retrospective (Multimedia Appendix 10) and prospective (Figure 4) cohorts, and more frequently (Figure 3, lower panel) over the future 6-month period.

**Figure 4.** Active case finding algorithm effectively risk-stratified the prospective patient cohort for future 6-month emergency department (ED) visit (graphic representation of low, medium, and high risk patients' time to next impending ED visit).



## Unsupervised Clustering: Subgroup Analysis of High-Risk Patients

Our principle component analysis retrospectively identified (Multimedia Appendix 11, left panel) and prospectively confirmed (Multimedia Appendix 11, right panel) a pattern of six distinct subgroups among the high-risk patients with risk scores greater than 70. These six clinically relevant clusters (retrospective: Multimedia Appendix 12, prospective: Multimedia Appendix 12) grouped around multiple "anchoring" demographic and chronic disease conditions. The chronic conditions co-occurred in many instances and included endocrine, nutritional, and metabolic diseases as well as immune disorders (ranging from 23.83%, 245/1028 to 74.21%, 590/795), diseases of the circulatory system (ranging from 13.7%, 99/722 to 68.4%, 544/795), diseases of the nervous system and sense organs (ranging from 26.0%, 188/722 to 66.5%, 529/795), diseases of the respiratory system (ranging from 23.44%, 241/1028 to 50.6%, 402/795), and diseases of the digestive system (ranging from 17.41%, 179/1028 to 55.0%, 437/795). These conditions were prevalent in all clusters, indicating that endocrine, immune, cardiac, nervous, respiratory, and digestive system dysfunctions co-occur. The largest cluster (#1) was characterized by predominantly adult female patients (between the ages of 19 and 49) characterized by chronic conditions including endocrine, nutritional, metabolic, and immune disorders, diseases of the sense organs, nervous, digestive, and respiratory systems. Cluster #6 was revealed as a high resource-consuming subgroup with the largest number of distinct chronic disease diagnoses accompanied by the largest number of laboratory and radiographic tests. In contrast, Cluster #5 contained a relatively younger population (age 19 to 34) with diminished incidence of chronic disease and minimal resource consumption. Clusters #2, 3, and 4 shared similar age, gender, and chronic disease distributions; however, these clusters displayed different usage profiles for tests and medication prescriptions. Moreover, in Clusters #2 and #3, approximately 20% of patients did not have any chronic disease diagnosis in the prior 12 months. Among Clusters #1 through #4, there are no direct correlations between the number of distinct chronic disease diagnoses and the usage of tests and prescriptions.

## Population Explorer Service: Statewide Real-Time Surveillance of Population Emergency Department Risks

Our ED predictive analytics were integrated into the Maine State HIE system (Figures 5 and 6) to allow real-time surveillance of population ED risks. Triggered by real-time iterative data feeds, each patient's ED risk measure can be updated periodically upon new data feed. This allows for trending risk scores over time, whereby targeting patients with major increases in risk may be as useful as targeting the patients at the highest risk. This Web-based population risk surveillance dashboard (Figure 6) empowers the ACO field staff and population health managers to visualize the ED risks derived from each resident's historical medical records in Maine. With our prospectively validated ED risk case finding algorithm, our coherent view of population ED risks can thus be feasible to resolve the barrier of the fragmented nature of population health information to improve public health practice and reduce ED utilization.

**Figure 5.** Schematic demonstration of data flow and communications of a population emergency department (ED) risk exploration system that allows real-time assessment of population ED risk.

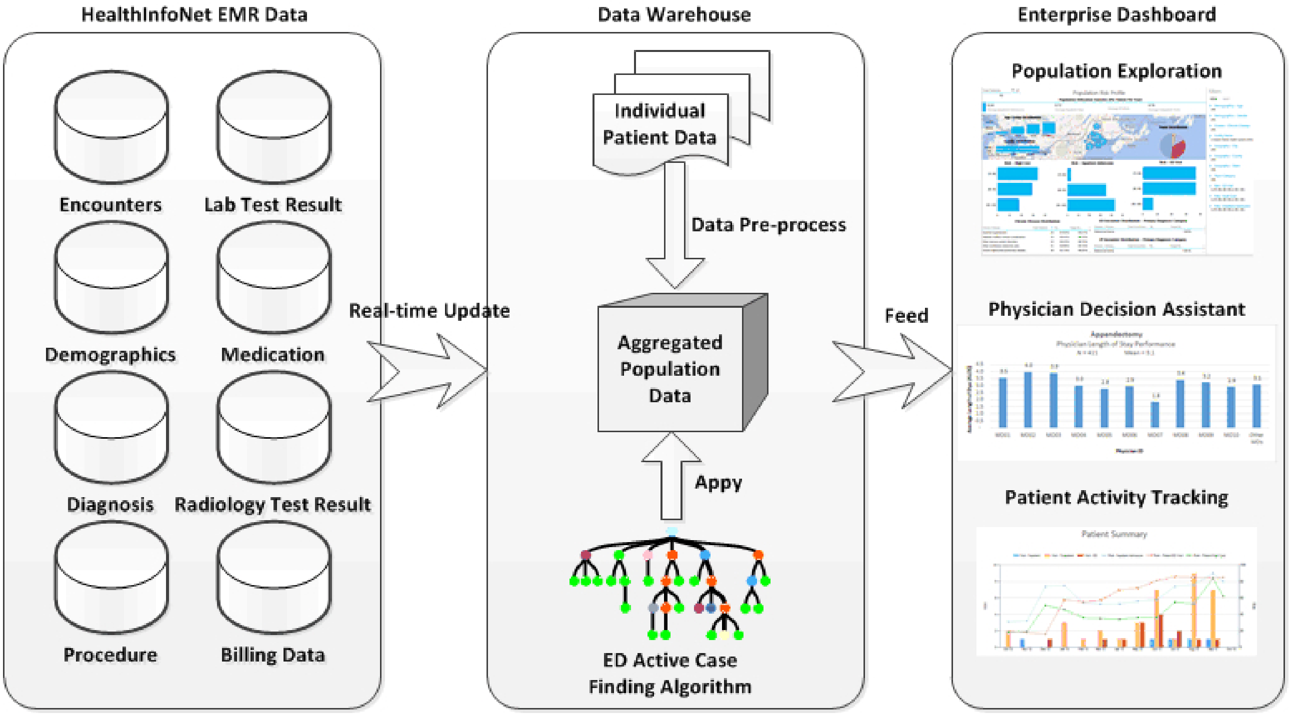
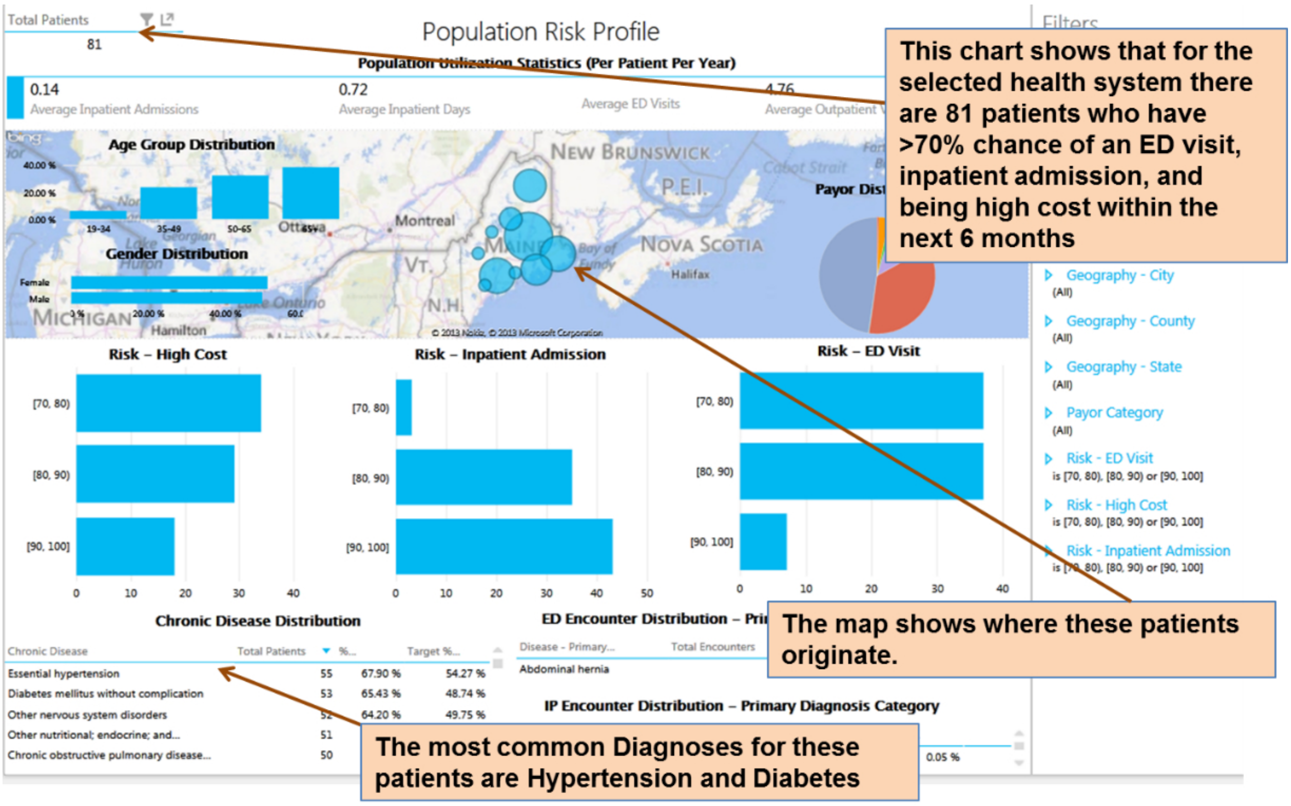


**Figure 6.** Total population emergency department (ED) risk monitoring dashboard.



| Chronic Disease | Total Patients | %... | Target %... |
|---|---|---|---|
| Essential hypertension | 55 | 67.90 % | 54.27 % |
| Diabetes mellitus without complication | 53 | 65.43 % | 48.74 % |
| Other nervous system disorders | 52 | 64.20 % | 49.75 % |
| Other nutritional; endocrine; and... | 51 | | |
| Chronic obstructive pulmonary disease... | 50 | | |

This chart shows that for the selected health system there are 81 patients who have >70% chance of an ED visit, inpatient admission, and being high cost within the next 6 months

The map shows where these patients originate.

The most common Diagnoses for these patients are Hypertension and Diabetes

## Discussion

### Principal Findings

We hypothesized that an individual patient's future 6-month ED risk can be forecast from the statistical learning of a population's comprehensive longitudinal clinical histories. Our use of the population-based HIE facilitated the development and prospective testing of the case finding algorithm presented here, which is population-based and not event-triggered (ED visit) analytics. After calculating the total population risk scores for future ED visit risk scores, this information is then made available to clinicians and caregivers at the point of care to support both individual patient and population-based decision making. Using adjustable risk settings allows multiple patient cohorts of different impending ED risks to be constructed. Moreover, high-risk patients with similar longitudinal clinical patterns can be subgrouped for targeted intervention in real time. Accurate identification of patient populations at high risk for ED visits is an integral component to address specific gaps in health care coverage, institute primary care-based interventions, and avoid preventable ED visits. Such active case finding may help providers deliver more efficient and effective health care interventions.

## Strengths and Limitations

Designed to be used in real time by population panel managers to forecast a future ED visit, our EMR-based active case finding method was prospectively validated with a reasonable level of sensitivity and specificity. To the best of our knowledge, our EMR-based population ED risk study is the first with such scale for ED trending across all payers, all diseases, and all age groups. Our study's obvious strength is the use of an entire US state in regard to predictive analytics. Its weakness is the study cohort's potential patient characteristics unique to the state of Maine, which may limit our model's general applicability to the other state populations.

Data limitations, for example, missing data, inaccurate diagnostic/procedure coding, and the unreliable tracking method to identify patients who die, may result in false negative and false positive case calls. Additionally, new patients who lack encounter histories tended to be categorized as low risk for future ED visits, a function that likely underestimates the ED risk for these subjects. We speculate that using additional currently non-reported features, including self-rated health conditions, lifestyle-related factors, and socioeconomic status may enhance the analytical approach to ED visit risk assessment.

Beyond identifying at-risk populations for potentially preventive services, gaining a deeper understanding of both the unique and common attributes of various subgroups may further facilitate overall management and the prevention of unwanted ED utilization. Moreover, to be clinically useful, a case finding model should be iterative and facilitate exploration of the potential benefit (PPV) or burden (false positive rate) (business case) of managing subpopulations of high-risk patients. Accordingly, we sought to determine whether unique patterns of resource utilization or clusters of patient subpopulations existed among the considerable heterogeneity of the high-risk patient population when considered together. We demonstrated that among the high-risk group patients, their associated demographics, chronic conditions, and varying patterns of resource consumption do not occur in isolation. Cluster analysis revealed six clinically relevant subgroups among the high-risk patient population that were confirmed as durable upon prospective testing. These subgroups have unique patterns of demographics, disease severities, comorbidities, and resource consumption, suggesting new opportunities to provide stratified care management among these groups. For example, Cluster #6 had senior patients with co-occurring histories of the most diverse chronic conditions and linked to the highest utilization of clinical tests and prescriptions. In addition, this group of patients is at considerable risk to experience poor health outcomes, including, but not limited to, lower quality of life, diminished functional status, as well as excess morbidity and mortality. This distinctive cluster could be targeted with new, enhanced care management strategies. We noted a decreased prevalence of the co-occurring chronic conditions in four other cluster groups of relatively younger adults with much less resource consumption. Within these four clusters, females aged 19-49 years without any chronic disease may benefit from targeted care to keep them out of the emergency room, although more analysis is needed to understand the risk drivers within this group. Currently, many existing care management strategies are directed toward single conditions and are event-triggered, for example, ED visit or hospital discharge. The current active case finding model provides novel opportunities to experiment with new strategies of coordinated care targeting a combination of conditions across different age and demographic groups that we speculate may lead to greater case management efficacy.

While the clusters identified in the study represent clinically similar populations that could guide specific care management strategies, we understand that the missing information (eg, mental health and substance abuse diagnostic information) may mask important characteristics of these clusters. Past studies have shown that mental health diagnoses are frequent within the ED patient population [23]. With data quality improving over time, we see a future opportunity for overall improvement in the predictive model and subsequent patient clustering.

With our ED risk model, tactics for modifying care management programs can be driven and measured against the analytical risk assessment derived from the HIE records. HIEs are a valuable data resource, providing longitudinal and comprehensive patient data. HIEs, such as HIN, that have completed the necessary rigorous mapping of multiple providers' data to standard nomenclature including LOINC [24], RXNorm [25], and SNOMED [26] offer an unparalleled data repository that can be leveraged to realize value through the application of advanced analytic techniques. However, while HIE data represents an ideal source of community-wide/regional patient data, operational HIEs are not present in all states. The predictive model and patient clustering method can be applied to any clinical dataset including the clinical EHRs directly as well as private HIEs within hospital networks.

## Conclusions

Our study is the first study of total population risk for ED visits across all payers, all diseases, and all age groups. Applying analytical tools on EMR and HIE data, including the active case finding model, the high-risk patient clustering method, and the Web-based real-time ED risk profiling analysis and exploration, will help health care providers effectively leverage their EMR to better understand ED service delivery while providing opportunities for improved health care delivery for the patients. A great strength of this work is the use of data from an entire state HIE, including data from across the entire spectrum of the health care system. This is not just hospital or emergency department data because it includes outpatient clinics and physician practices. In that regard, our work should serve as a model of what other states can do with HIE data to really impact patient care and population health.

## Acknowledgments

We express our gratitude to the hospitals, medical practices, physicians, and nurses participating in Maine's HIE. We also thank the biostatistics colleagues at the Department of Health Research and Policy, Stanford Pediatric Proteomics Group for critical discussions.

## Conflicts of Interest

KGS, EW, and XBL are co-founders and equity holders of HBI Solutions, Inc., which is currently developing predictive analytics solutions in health care. From the Departments of Pediatrics, Surgery, and Statistics, Stanford University School of Medicine, Stanford, California, AYS, SH, ZL, YW, KGS, XBL conducted this research as part of a personal outside consulting arrangement with HBI Solutions, Inc. The research and research results are not, in any way, associated with Stanford University.

## Multimedia Appendix 1

Electronic medical record (EMR) features used to develop active case finding model.

[PDF File (Adobe PDF File), 9KB - ijmr_v4i1e2_app1.pdf ]

## Multimedia Appendix 2

Emergency department (ED) admission "time-to-event curve" showed pattern of rapid accrual with stable and consistent ED visit rate thereafter. Population ED visit curves, of patients with more than one or any ED visit, stabilized within 6 months from evaluation time, indicating a 6-month cutoff is clinically reasonable. Assessment date: January 1, 2013.

[PNG File, 28KB - ijmr_v4i1e2_app2.png ]

## Multimedia Appendix 3

Study cohort construction, and inclusion/ exclusion criteria; retrospective/ prospective cohort construction.

[PDF File (Adobe PDF File), 94KB - ijmr_v4i1e2_app3.pdf ]

## Multimedia Appendix 4

Patient characteristics.

[PDF File (Adobe PDF File), 10KB - ijmr_v4i1e2_app4.pdf ]

## Multimedia Appendix 5

Exploratory data analysis: patient counts of total set and those having emergency department (ED) revisit in future 6 months, as function of number of chronic diagnoses (left panel) and ED visits in past 12 months (right panel), and percentages of patients with ED revisits was also plotted.

[PNG File, 52KB - ijmr_v4i1e2_app5.png ]

## Multimedia Appendix 6

Technical details of decision tree based modeling.

[PDF File (Adobe PDF File), 112KB - ijmr_v4i1e2_app6.pdf ]

## Multimedia Appendix 7

Feature selection and characterization of discriminant features in retrospective/prospective dataset.

[PDF File (Adobe PDF File), 156KB - ijmr_v4i1e2_app7.pdf ]

## Multimedia Appendix 8

The model performance was gauged by ROC analysis for retrospective blind testing and perspective validating respectively.

[PNG File, 59KB - ijmr_v4i1e2_app8.png ]

## Multimedia Appendix 9

Unsupervised clustering of high risk population using principal component analysis.

[PDF File (Adobe PDF File), 10KB - ijmr_v4i1e2_app9.pdf ]

---

## Multimedia Appendix 10

Active case finding algorithm effectively risk-stratified retrospective patient cohort for future 6-month emergency department (ED) visit: graphic representation of low-, medium-, and high-risk patients' time to the next impending ED visit.

[PNG File, 28KB - ijmr_v4i1e2_app10.png ]

---

## Multimedia Appendix 11

Unsupervised clustering of the high-risk patients identified consistent distinct subgroups in both retrospective (left panel) and prospective (right panel) cohorts.

[PNG File, 351KB - ijmr_v4i1e2_app11.png ]

---

## Multimedia Appendix 12

Clustering of emergency department 6-month high-risk patients in the retrospective/prospective cohort according to demographics and prior year clinical histories.

[PDF File (Adobe PDF File), 29KB - ijmr_v4i1e2_app12.pdf ]

---

## References

1. Bottle A, Aylin P, Majeed A. Identifying patients at high risk of emergency hospital admissions: a logistic regression analysis. J R Soc Med 2006 Aug;99(8):406-414 [FREE Full text] [doi: 10.1258/jrsm.99.8.406] [Medline: 16893941]
2. McDonald AJ, Pelletier AJ, Ellinor PT, Camargo CA. Increasing US emergency department visit rates and subsequent hospital admissions for atrial fibrillation from 1993 to 2004. Ann Emerg Med 2008 Jan;51(1):58-65. [doi: 10.1016/j.annemergmed.2007.03.007] [Medline: 17466409]
3. Pines JM, Mutter RL, Zocchi MS. Variation in emergency department admission rates across the United States. Med Care Res Rev 2013 Apr;70(2):218-231. [doi: 10.1177/1077558712470565] [Medline: 23295438]
4. Kharbanda AB, Hall M, Shah SS, Freedman SB, Mistry RD, Macias CG, et al. Variation in resource utilization across a national sample of pediatric emergency departments. J Pediatr 2013 Jul;163(1):230-236. [doi: 10.1016/j.jpeds.2012.12.013] [Medline: 23332463]
5. Taubman SL, Allen HL, Wright BJ, Baicker K, Finkelstein AN. Medicaid increases emergency-department use: evidence from Oregon's Health Insurance Experiment. Science 2014 Jan 17;343(6168):263-268. [doi: 10.1126/science.1246183] [Medline: 24385603]
6. Cutler D. Analysis & commentary. How health care reform must bend the cost curve. Health Aff (Millwood) 2010 Jun;29(6):1131-1135 [FREE Full text] [doi: 10.1377/hlthaff.2010.0416] [Medline: 20530342]
7. Billings J, Dixon J, Mijanovich T, Wennberg D. Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. BMJ 2006 Aug 12;333(7563):327 [FREE Full text] [doi: 10.1136/bmj.38870.657917.AE] [Medline: 16815882]
8. Kansagara D, Englander H, Salanitro A, Kagen D, Theobald C, Freeman M, et al. Risk prediction models for hospital readmission: a systematic review. JAMA 2011 Oct 19;306(15):1688-1698 [FREE Full text] [doi: 10.1001/jama.2011.1515] [Medline: 22009101]
9. Hustey FM, Mion LC, Connor JT, Emerman CL, Campbell J, Palmer RM. A brief risk stratification tool to predict functional decline in older adults discharged from emergency departments. J Am Geriatr Soc 2007 Aug;55(8):1269-1274. [doi: 10.1111/j.1532-5415.2007.01272.x] [Medline: 17661968]
10. Meldon SW, Mion LC, Palmer RM, Drew BL, Connor JT, Lewicki LJ, et al. A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department. Acad Emerg Med 2003 Mar;10(3):224-232. [Medline: 12615588]
11. Mion LC, Palmer RM, Meldon SW, Bass DM, Singer ME, Payne SM, et al. Case finding and referral model for emergency department elders: a randomized clinical trial. Ann Emerg Med 2003 Jan;41(1):57-68. [doi: 10.1067/mem.2003.3] [Medline: 12514683]
12. Shelton P, Sager MA, Schraeder C. The community assessment risk screen (CARS): identifying elderly persons at risk for hospitalization or emergency department visit. Am J Manag Care 2000 Aug;6(8):925-933 [FREE Full text] [Medline: 11186504]

13. Sin DD, Man SF. Low-dose inhaled corticosteroid therapy and risk of emergency department visits for asthma. Arch Intern Med 2002 Jul 22;162(14):1591-1595. [Medline: 12123402]
14. Hao S, Jin B, Shin AY, Zhao Y, Zhu C, Li Z, et al. Risk prediction of emergency department revisit 30 days post discharge: a prospective study. PLoS One 2014;9(11):e112944 [FREE Full text] [doi: 10.1371/journal.pone.0112944] [Medline: 25393305]
15. He X, Ji M, Zhang C, Bao H. A variance minimization criterion to feature selection using laplacian regularization. IEEE Trans Pattern Anal Mach Intell 2011 Mar 1. [doi: 10.1109/TPAMI.2011.44] [Medline: 21383399]
16. Healthcare Cost and Utilization Project (HCUP). Chronic condition indicator (CCI). 2014. URL: http://www.hcup-us.ahrq.gov/toolssoftware/chronic/chronic.jsp [accessed 2014-11-11] [WebCite Cache ID 6TzpZJUlw]
17. Breiman L. Bagging predictors. Machine Learning 1996;24(2):123-140. [doi: 10.1023/A:1018054314350]
18. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. Ann Appl Statist 2008;2(3):841-860. [doi: 10.1023/A:1018054314350]
19. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002 May 14;99(10):6567-6572 [FREE Full text] [doi: 10.1073/pnas.082099299] [Medline: 12011421]
20. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005 Oct 15;21(20):3940-3941 [FREE Full text] [doi: 10.1093/bioinformatics/bti623] [Medline: 16096348]
21. Jolliffe IT. Principal component analysis. New York: Springer; 2002.
22. Zha H, He X, Ding C, Gu M, Simon HD. Spectral relaxation for k-means clustering. In: Advances in Neural Information Processing Systems 14. 2001 Presented at: NIPS 2001; 2001; Vancouver, Canada p. 1057-1064.
23. Olfson M, Marcus SC, Bridge JA. Emergency department recognition of mental disorders and short-term outcome of deliberate self-harm. Am J Psychiatry 2013 Dec;170(12):1442-1450. [doi: 10.1176/appi.ajp.2013.12121506] [Medline: 23897218]
24. Forrey AW, McDonald CJ, DeMoor G, Huff SM, Leavelle D, Leland D, et al. Logical observation identifier names and codes (LOINC) database: a public use set of codes and names for electronic reporting of clinical laboratory test results. Clin Chem 1996 Jan;42(1):81-90 [FREE Full text] [Medline: 8565239]
25. Bennett CC. Utilizing RxNorm to support practical computing applications: capturing medication history in live electronic health records. J Biomed Inform 2012 Aug;45(4):634-641 [FREE Full text] [doi: 10.1016/j.jbi.2012.02.011] [Medline: 22426081]
26. Wingert F. An indexing system for SNOMED. Methods Inf Med 1986 Jan;25(1):22-30. [Medline: 3753739]

## Abbreviations

**ACO:** accountable care organization
**AUC:** area under the curve
**ED:** emergency department
**EMR:** electronic medical record
**HIE:** health information exchange
**HIN:** HealthInfoNet
**ICD-9-CM:** International Classification of Diseases, 9th Revision, Clinical Modification
**PPV:** positive predictive value
**ROC:** receiver operating characteristic

*Edited by G Eysenbach; submitted 12.11.14; peer-reviewed by H Wu; comments to author 02.12.14; revised version received 06.12.14; accepted 11.12.14; published 13.01.15.*

*Please cite as:*
*Hu Z, Jin B, Shin AY, Zhu C, Zhao Y, Hao S, Zheng L, Fu C, Wen Q, Ji J, Li Z, Wang Y, Zheng X, Dai D, Culver DS, Alfreds ST, Rogow T, Stearns F, Sylvester KG, Widen E, Ling XB*
*Real-Time Web-Based Assessment of Total Population Risk of Future Emergency Department Utilization: Statewide Prospective Active Case Finding Study*
*Interact J Med Res 2015;4(1):e2*
*URL: http://www.i-jmr.org/2015/1/e2/*
*doi:10.2196/ijmr.4022*
*PMID:25586600*

Review

# Information Seeking in Social Media: A Review of YouTube for Sedentary Behavior Content

Emily Knight[1*], BSc; Brittany Intzandt[2*], BSc; Alicia MacDougall[3*]; Travis J Saunders[4*], BSc, MSc, PhD

[1]University of Western Ontario, Health and Rehabilitation Sciences, London, ON, Canada
[2]Wilfrid Laurier University, Waterloo, ON, Canada
[3]University of Western Ontario, Kinesiology, London, ON, Canada
[4]University of Prince Edward Island, Faculty of Science, Charlottetown, PE, Canada
[*]all authors contributed equally

**Corresponding Author:**
Emily Knight, BSc
University of Western Ontario
Health and Rehabilitation Sciences
1201 Western Road
London, ON, N6G 1H1
Canada
Phone: 1 519 685 4292
Fax: 1 519 685 4071
Email: eknigh2@uwo.ca

## Abstract

**Background:** The global prevalence of sedentary lifestyles is of grave concern for public health around the world. Moreover, the health risk of sedentary behaviors is of growing interest for researchers, clinicians, and the general public as evidence demonstrates that prolonged amounts of sedentary time increases risk for lifestyle-related diseases. There is a growing trend in the literature that reports how social media can facilitate knowledge sharing and collaboration. Social sites like YouTube facilitate the sharing of media content between users.

**Objective:** The purpose of this project was to identify sedentary behavior content on YouTube and describe features of this content that may impact the effectiveness of YouTube for knowledge translation.

**Methods:** YouTube was searched on a single day by 3 independent reviewers for evidence-based sedentary behavior content. Subjective data (eg, video purpose, source, and activity type portrayed) and objective data (eg, number of views, comments, shares, and length of the video) were collected from video.

**Results:** In total, 106 videos met inclusion criteria. Videos were uploaded from 13 countries around the globe (ie, Australia, Barbados, Belgium, Canada, Colombia, Kenya, New Zealand, Russia, South Africa, Spain, Ukraine, United Kingdom, United States). The median video length was 3:00 minutes: interquartile range (IQR) 1:44-5:40. On average, videos had been on YouTube for 15.0 months (IQR 6.0-27.5) and had been viewed 239.0 times (IQR 44.5-917.5). Videos had remarkably low numbers of shares (median 0) and comments (median 1). Only 37.7% (40/106) of videos portrayed content on sedentary behaviors, while the remaining 66 videos portrayed physical activity or a mix of behaviors. Academic/health organizations (39.6%, 42/106) and individuals (38.7%, 41/106) were the most prevalent source of videos, and most videos (67.0%, 71/106) aimed to educate viewers about the topic.

**Conclusions:** This study explored sedentary behavior content available on YouTube. Findings demonstrate that there is confusion between physical activity and sedentary behaviors, that content is being uploaded to the site from around the globe, that content is primarily from health organizations and individuals with the purpose of educating fellow users, but that low views, comments, and shares suggest that sedentary behavior content is relatively underutilized on YouTube. Future research may wish to leverage social platforms, such as YouTube, to facilitate implementation and sharing of evidence-based sedentary behavior content.

***(Interact J Med Res 2015;4(1):e3)*** doi:10.2196/ijmr.3835

**KEYWORDS**

sedentary lifestyle; translational medical research; social media

## Introduction

### Sedentary Behavior

Since the 1950s, researchers have understood the importance of physical activity in promoting health [1]. An active lifestyle is now understood to reduce both the morbidity and mortality of a wide range of chronic diseases ranging from cancer to cardiovascular disease [2]. Insufficient levels of physical activity are responsible for 6% of global mortality and are the fourth leading cause of death around the world [3].

While the health importance of physical activity is well established, recent evidence suggests that sedentary behavior (eg, activities done while sitting) [4] also plays an important role in the development of chronic disease. For example, sedentary behaviors are associated with adverse health effects, including undesirable changes in cardiometabolic markers, vascular, bone, and psychosocial health independent of exercise [5-8]. This is important, given that it is possible to accumulate high amounts of both sedentary and exercise behaviors during a single day [6,7,9-11]. Further, the proportion of time spent in sedentary behavior dwarfs that spent in physical activity. For example, Canadian adults spend roughly 70% of their waking hours engaging in sedentary behavior, while just 3% engaged in moderate or vigorous physical activity [12]. Not surprisingly, recent studies have estimated that sedentary behavior may reduce the life expectancy of Western nations by 1-2 years [13,14]. Given these important differences, researchers have recently argued that sedentary behavior should be viewed as an independent and distinct construct, rather than simply the lack of physical activity [4]. Further, it has been shown that sedentary behavior and physical activity have independent and distinct relationships with health.

### Knowledge Translation

The knowledge-to-action framework from Graham et al [15] highlights the multifaceted nature of knowledge creation and implementation. Knowledge about the health effects of prolonged sedentary behaviors has been generated (ie, the knowledge creation cycle). However, a gap exists in understanding how this information is being implemented to impact the health of lay users (ie, the knowledge utilization cycle).

In North America, the Internet is a primary source of health information, with more than half of users seeking health information online [16-18]. Previous evidence has reported that over half of American and European citizens have used the Internet to seek health-related information [19]. Online mediums may help to bridge the research to action gap by allowing evidence-based information to freely reach the homes of a broad spectrum of users. There is a growing trend in the literature reporting how social media can facilitate knowledge sharing and collaboration [20].

### YouTube

YouTube is among the top three most popular websites visited around the world, with more than 4 billion videos being watched by users daily [21,22]. The primary concept of social media sites like YouTube is the sharing of media content between users [23]. Videos allow for the sharing of complex ideas in a simple format [22]. The Health Care Social Media List form the Mayo Clinic identifies over 700 health-related associations in the United States that have established a presence on YouTube [24]. A strength of the social media format for knowledge translation is its capacity for timely updates, in contrast to the slow uptake and evolution of information shared through traditional peer-reviewed formats [19]. However, there is a lack of regulation on the content available through YouTube. Previous research has reported that misleading information posted through videos on YouTube could endanger viewers [22].

### Purpose

There is a growing body of evidence exploring health content available through YouTube. Research has identified that health-related videos posted to YouTube may contain erroneous and potentially harmful health information [20,22,25]. As knowledge about the health risk of sedentary behavior transitions from research to practice, there is value in understanding if the information available on YouTube is evidence-based. Therefore, the purpose of this project was to identify sedentary behavior content on YouTube and describe features of this content that may impact the effectiveness of YouTube for knowledge translation, such as evidence-informed messages and description of video characteristics. Understanding what information is currently available may help researchers tailor their messages to promote more effective knowledge translation and uptake.

## Methods

### Search Strategy

YouTube was searched on a single day (May 25, 2014) by 3 independent reviewers (EK, BI, AM). Computers were set to “incognito”/“worldwide” to limit any filtering by the site to previous user data and to help ensure that search results would not be limited to local country of searching but instead include videos from around the globe.

Similar to the methods of Williams et al [26], Google AdWords [27] was used to search keyword phrases frequently used by the public when searching Google in relation to “sedentary behavior” to develop a list of popular phrases for the topic that could be searched on YouTube. Subsequently, three keyword searches of YouTube were conducted (“sedentary behaviour”, “sedentary behavior”, and “sedentary lifestyle”) as well as a category search of the YouTube-generated “sedentary lifestyle” topic channel. Currently, YouTube algorithms create topic channels based on volume of content on the site, in the present case linked to the keyword phrase “sedentary lifestyle”. The channel’s main page presents three popular videos sorted by relevance to the topic channel for six subcategories: (1) popular sedentary lifestyle videos, (2) popular sedentary lifestyle and health videos, (3) popular sedentary lifestyle and physical exercise videos, (4) popular sedentary lifestyle and obesity videos, (5) popular sedentary lifestyle & lifestyle videos, and (6) popular sedentary lifestyle and childhood obesity videos.

For each of the four searches (three keyword, one topic channel), results were sorted by both relevance and views using the

YouTube search features, and the first two result pages (approximately 40 videos) were assessed. The goal of the search process was to identify YouTube content that users are accessing the most. Searching was limited to the initial two pages of results based on the following a priori criteria: (1) saturation of the topic becomes evident in most cases by the end of the second page at which point search results are obviously not related to the search phrase, and unlike the searching process employed in systematic reviews, (2) typical YouTube users may be less likely to continue scanning results past the initial two pages of results, especially when videos appear not to be linked to the search phrase. Videos were excluded if they were not available in English or portrayed content obviously not related to the search (eg, music videos, product advertisements). Additionally, individual users' channels that were generated in the search results were not reviewed, as they contained multiple videos posted by the user with varying relation to the desired search content.

### Data Management and Analysis

Video data were coded into an electronic spreadsheet and analyzed in June 2014. For each video, we collected both objective data (video title, URL, number of views, number of shares, the length of the video, number of comments posted by YouTube users, descriptive text and keywords that the user who uploaded the video included, and the YouTube category used to classify the video) and subjective data (the purpose and source of the video, and the type of activity content included in the video). Textbox 1 defines coding themes used for subjective data. To ensure consistency in coding, a representative video from each search was coded collaboratively until consensus was reached.

Descriptive statistics were performed to understand the context of evidence available on YouTube. Specifically, there was interest in understanding if the information available on YouTube represents best evidence for sedentary behavior, if users are seeking information on sedentary behavior through this medium, and describing video content in terms of who is producing the video to help inform future initiatives for leveraging YouTube as a knowledge translation vehicle. To explore the descriptive statistics by popularity of content on the YouTube site, interquartile range of view counts were used to group videos based on number of views. Pearson correlation coefficient (*r*) was calculated to explore the relationship between views, keyword search, video length, video source, and activity classification. Statistical significance was set at $P<.05$.

**Textbox 1.** Definition of coding themes for subjective data extraction.

Video Source

- Media: Video presented by an identified news/media source
- Individual: Video presented by an individual
- Academic/Health Organization: Video presented by an academic conference, research group, or medical organization
- Consumer: Video endorsing and/or promoting sale of a product/service

Video Purpose

- Educate: Video informs/teaches about the topic, which includes evidence-based information
- Opinion: Video portraying an individual's or organization's perspective on the topic
- Academic Presentation: Video of a presentation to academic audiences (eg, conference proceedings)
- Commercial: Video promoting a company's or individual's product(s)

Activity Classification

- Physical Activity: Video portraying information on physical activity and the health benefits and/or public health recommendations for this activity
- Sedentary Behavior: Video portraying information on the behavior and/or health outcomes of activities in a sitting or reclining posture
- Mixed: Video incorporating information of both physical activity as well as activities in sitting/reclining postures

## Results

Figure 1 shows the search results. The initial search yielded 232 videos. After removing ineligible and duplicate results, data were analyzed from 106 videos. Table 1 summarizes the included videos. The location of the uploader was not discernable via the YouTube site. Therefore, an independent online database (YouTube Stats [28]) was used to search the country of origin for videos. It should be noted that the database provides only a source of origin for videos with >4 subscribers; therefore, it did not provide a complete source for origin of all videos in this sample (Table 2). Figure 2 shows the distribution of videos around the globe and demonstrates a substantial representation of content from North America.

**Table 1.** Video results.

| Result No. | URL | Title | Months on YouTube | Length (min) | Views |
|---|---|---|---|---|---|
| 1 | http://www.youtube.com/watch?v=dJY9NWoA3Dk | 10 Minute Basic Workout for the Sedentary Individual | 19 | 3:40 | 495 |
| 2 | https://www.youtube.com/watch?v=o12kXL0iopE | 10-1. Sedentary Lifestyle and How To Improve Cardiovascular Endurance with Exercise | 61 | 6:39 | 13,100 |
| 3 | http://www.youtube.com/watch?v=oWW-Ws32MLI | 15. Osteoporosis Sedentary Lifestyle | 10 | 1:30 | 19 |
| 4 | http://www.youtube.com/watch?v=YaucXroi8ls | 2012 JustStand Wellness Summit; Dr David Dunstan, Baker IDI Heart and Diabetes Institute | 9 | 59:41 | 67 |
| 5 | http://www.youtube.com/watch?v=lJG4T5LpDzM | 2014 Physical Activity Forum - Get Up, Stand Up | 0 | 1:03:03 | 12 |
| 6 | https://www.youtube.com/watch?v=6uomlJh5g9g | 8 Weight Loss Tips for a Sedentary Lifestyle | 23 | 3:16 | 65 |
| 7 | http://www.youtube.com/watch?v=aoMmpFbZz-k | American Idle: Sedentary Time and Health | 1 | 3:28 | 0 |
| 8 | https://www.youtube.com/watch?v=MWVv0Z3x_xA | America's Walking \| A Call to Action \| A Sedentary Lifestyle | 50 | 8:54 | 913 |
| 9 | https://www.youtube.com/watch?v=3WxP7fU_JF4 | An Introduction to Active vs. Sedentary Lifestyles as it Relates to Chronic Disease.wmv | 50 | 2:05 | 3822 |
| 10 | http://www.youtube.com/watch?v=sAo_352QYjs | Are you Active or Sedentary? How can you become more Active for FREE? Easy tips Lisa in Marbella | 2 | 4:30 | 54 |
| 11 | https://www.youtube.com/watch?v=NndqoguNrzs | Avoid Sedentary Lifestyle | 2 | 6:40 | 14 |
| 12 | http://www.youtube.com/watch?v=hMegW5G7ZEk | Back Pain & Sedentary Life Style | 19 | 4:58 | 258 |
| 13 | http://www.youtube.com/watch?v=tnQ1Ye6J5Aw | Beating Sedentary Behavior at Prince of Wales school | 11 | 1:39 | 150 |
| 14 | http://www.youtube.com/watch?v=clZq8w2lRHs | Benefits of a Standing Desk on your Feet Australia Campaign | 3 | 5:25 | 90 |
| 15 | http://www.youtube.com/watch?v=apEYRbfVsks | Bonnie Spring: Can we Design our Way out of the Obesity Epidemic? | 31 | 42:39 | 132 |
| 16 | http://www.youtube.com/watch?v=63SMM8tTELw | Bouncing at Work | 9 | 0:47 | 101 |
| 17 | https://www.youtube.com/watch?v=B9jhqaXZR0w | Breaking the Sedentary Lifestyle | 10 | 2:52 | 17 |
| 18 | http://www.youtube.com/watch?v=uqviPmuytQA | Breaks in Sedentary Time are Associated with Reduced Health Risk in Children. | 5 | 4:18 | 128 |
| 19 | https://www.youtube.com/watch?v=5ty7GhKJ0Yg | Camp Abilities: A Vision of a Healthier Lifestyle | 46 | 3:08 | 1079 |
| 20 | http://www.youtube.com/watch?v=UFg4amY6ltg | Classroom & Sedentary Behavior | 6 | 6:49 | 53 |
| 21 | http://www.youtube.com/watch?v=CLSbS0yEJ5M | Classroom Teacher Challenges for Managing PA: Reduce Sedentary Behavior with Strctu | 21 | 1:21 | 922 |
| 22 | https://www.youtube.com/watch?v=d4g7bPS_8pk | Combatting a Sedentary Lifestyle - Penn State Hershey | 12 | 1:00 | 517 |
| 23 | http://www.youtube.com/watch?v=osnXMZg4ccQ | Course Director Pitch - BSc Hons Physical Activity Exercise and Health | 0 | 2:23 | 44 |
| 24 | http://www.youtube.com/watch?v=sSKF0RoAOUY | Dan Oliver and Ryan Durden's Video Presentation | 1 | 4:46 | 12 |
| 25 | http://www.youtube.com/watch?v=lyBKXKBei8o | Dangers of a Sedentary Lifestyle | 12 | 8:06 | 461 |
| 26 | https://www.youtube.com/watch?v=to8GmlDfhmw | Dangers of a Sedentary Lifestyle | 26 | 1:20 | 718 |
| 27 | http://www.youtube.com/watch?v=eHoVLMFboBI | DynaCubes Breaking Sedentary behavior | 11 | 1:36 | 32 |
| 28 | https://www.youtube.com/watch?v=YRJpfkqYBp4 | Easy Ways to Increase Physical Activity | 49 | 0:58 | 12,298 |
| 29 | https://www.youtube.com/watch?v=rasZGZpQsy0 | Educating the Student Body: Taking Physical Activity and Physical Education to School | 12 | 2:40 | 10,810 |

| Result No. | URL | Title | Months on YouTube | Length (min) | Views |
|---|---|---|---|---|---|
| 30 | http://www.youtube.com/watch?v=B7r6r9UOxu0 | Effect of Physical Activity on Serum Prostate-Specific Antigen Concentrations | 16 | 5:05 | 117 |
| 31 | http://www.youtube.com/watch?v=WwOdG1INtV8 | EPI-NPAM 2012- Sedentary Behavior, Physical Activity and Incident Coronary Heart Disease | 26 | 7:26 | 366 |
| 32 | https://www.youtube.com/watch?v=PZVkn-pDsoNY | Erin has been Overweight her Whole Life | 52 | 3:42 | 2003 |
| 33 | http://www.youtube.com/watch?v=Iz0JgVoEFHc | Exercise & Weight Loss - Episode 7 - Summer Tomato Live | 36 | 43:12:00 | 1343 |
| 34 | http://www.youtube.com/watch?v=QFc-5oXgbYY | Exercise training Alters Subcutaneous White Adipose Tissue (scWAT) | 11 | 2:18 | 684 |
| 35 | http://www.youtube.com/watch?v=1PvQjNF2ths | Fitness Paradigm | 21 | 43:25 | 56 |
| 36 | https://www.youtube.com/watch?v=IqWfxxux-mi4 | Give Up Your Sedentary Lifestyle | 37 | 2:37 | 594 |
| 37 | http://www.youtube.com/watch?v=Z_cx-n_7mXg | Gregory Norman - Physical Activity and Sedentary Behavior Classification Using Motion Sensor and SM | 7 | 24:15 | 38 |
| 38 | https://www.youtube.com/watch?v=Bxt0fplop-vA | Health & Fitness Tips For Truck Drivers Revealed By Twin Drivers | 4 | 1:08 | 1434 |
| 39 | https://www.youtube.com/watch?v=P5ve869jb_Y | Healthy Eating vs. Sedentary Lifestyle - Fabio Viviani | 12 | 4:55 | 1201 |
| 40 | http://www.youtube.com/watch?v=A7vQWf-miVs | HK200 Ken Etics | 2 | 2:06 | 1991 |
| 41 | https://www.youtube.com/watch?v=5Nj3smp-fUtM | How to Avoid a Sedentary Lifestyle: | 46 | 1:47 | 5964 |
| 42 | https://www.youtube.com/watch?v=PkpjnGHeNN4 | How to Pronounce Sedentary | 14 | 0:20 | 2151 |
| 43 | https://www.youtube.com/watch?v=aDRYEYSb_f8 | Informative Speech-Sedentary Lifestyle-Wilkey | 13 | 7:42 | 30 |
| 44 | http://www.youtube.com/watch?v=6BYvKdi-Wtcw | Intro.wmv | 0 | 2:41 | 13 |
| 45 | https://www.youtube.com/watch?v=ovAev4W7BeY | Is Korea Affected by an Abdominal Obesity and Sedentary Lifestyle Epidemic? | 15 | 0:37 | 45 |
| 46 | http://www.youtube.com/watch?v=4MnQ7XnBpcc | Jigsaw Desks | 49 | 2:23 | 119 |
| 47 | https://www.youtube.com/watch?v=2Zquq7I_Ol0 | Joe Rogan on Fresh Food and a Sedentary Lifestyle | 2 | 5:42 | 328 |
| 48 | http://www.youtube.com/watch?v=BcxBe-VaGnjI | Keynote Speech - Dr Kong Chen - Be Active 2012 | 18 | 50:16 | 147 |
| 49 | http://www.youtube.com/watch?v=Mbr7rDe7vRw | L3 Health Online task Intro.wmv | 30 | 2:20 | 27 |
| 50 | https://www.youtube.com/watch?v=TOT6T--70_w | Lesson 2, The Benefits of a Healthy Lifestyle | 30 | 2:10 | 741 |
| 51 | https://www.youtube.com/watch?v=oh40z8MOzh0 | Lifestyle: Ageing and Health | 25 | 3:43 | 1866 |
| 52 | http://www.youtube.com/watch?v=pH_iFV3nYnQ | Motivations for Continued Involvement in Physical Activity | 22 | 4:33 | 348 |
| 53 | http://www.youtube.com/watch?v=El-DA0YzORjs | Obesity: A Heavy Burden | 4 | 1:07:52 | 239 |
| 54 | http://www.youtube.com/watch?v=Ok96iSLWyyg | Older Women Spend 2/3rd of Time Sedentary | 3 | 0:33 | 10 |
| 55 | http://www.youtube.com/watch?v=3fV91ZGAkR0 | Physical Activity and Sedentary Jobs | 36 | 4:47 | 406 |
| 56 | http://www.youtube.com/watch?v=XTTCKd8pZNQ | Prof. Stuart Biddle - Teaser | 7 | 0:50 | 23 |
| 57 | http://www.youtube.com/watch?v=yLJo5VECe5E | Project Play: Reimagining Youth Sports in America | 3 | 1:14:19 | 4459 |
| 58 | http://www.youtube.com/watch?v=9wRDIXxB-mIY | Promoting More Physical Activity and Less Sedentary Behaviour in Young People | 7 | 0:27 | 214 |
| 59 | https://www.youtube.com/watch?v=L2E8_MJsNZA | PSA Sedentary Lifestyle | 24 | 5:35 | 513 |

| Result No. | URL | Title | Months on YouTube | Length (min) | Views |
|---|---|---|---|---|---|
| 60 | http://www.youtube.com/watch?v=hmCIuFv05ag | Scottish Kids Less Active | 0 | 2:34 | 2 |
| 61 | http://www.youtube.com/watch?v=ZuAKXAGZK0w | Sedentary Behavior | 25 | 7:22 | 37 |
| 62 | http://www.youtube.com/watch?v=qondXFSjPPA | Sedentary Behavior and your Health | 5 | 0:51 | 31 |
| 63 | http://www.youtube.com/watch?v=znXimY_iNvs | Sedentary Behavior in College Students | 10 | 0:55 | 28 |
| 64 | http://www.youtube.com/watch?v=Dr226ZCZuPw | Sedentary Behavior in Youth | 6 | 5:11 | 24 |
| 65 | http://www.youtube.com/watch?v=ynfUyMN6ReQ | Sedentary Behavior- Target for Change, Challenge to Assess | 25 | 13:17 | 1056 |
| 66 | http://www.youtube.com/watch?v=qondXFSjPPA | Sedentary Behaviour & Health: Is the Chair the Most Important Threat to Health in 21st Century? | 0 | 31:56 | 4 |
| 67 | http://www.youtube.com/watch?v=rTRHkUuLON8 | Sedentary Behaviour (Get off the couch) | 14 | 2:34 | 233 |
| 68 | http://www.youtube.com/watch?v=hWWHhvxYzXk | Sedentary Behaviour Researchers - A Guaranteed Standing Ovation | 19 | 0:10 | 179 |
| 69 | http://www.youtube.com/watch?v=gB33PRJttyU | Sedentary Behaviour: Not Even Once | 14 | 2:17 | 252 |
| 70 | http://www.youtube.com/watch?v=qondXFSjPPA | Sedentary Life—Barriers to Physical Activities | 26 | 23:36 | 644 |
| 71 | http://www.youtube.com/watch?v=8_xeukXJPbk | Sedentary Lifestyle | 16 | 4:15 | 56 |
| 72 | https://www.youtube.com/watch?v=5b064VsRiiY | Sedentary Lifestyle | 1 | 1:45 | 81 |
| 73 | https://www.youtube.com/watch?v=2lWnUnQTfxU | Sedentary lifestyle | 12 | 4:17 | 17 |
| 74 | https://www.youtube.com/watch?v=cWaVOq0AAAQ | Sedentary Lifestyle - Get Fit or Get Fat | 36 | 2:29 | 1451 |
| 75 | https://www.youtube.com/watch?v=F2PP7vDJh0 | Sedentary Lifestyle - It's Bad!!! This's How I solved My Problem of Chronic Sedentary Lifestyle - p1 | 27 | 10:39 | 244 |
| 76 | https://www.youtube.com/watch?v=oVGi6jZ99ys | Sedentary Lifestyle and Obesity | 20 | 2:20 | 3738 |
| 77 | http://www.youtube.com/watch?v=-Y68CfN6oCk | Sedentary Lifestyle As Damaging As Smoking, Study Says | 46 | 1:55 | 1651 |
| 78 | https://www.youtube.com/watch?v=tupL_3uAZx0 | Sedentary Lifestyle Doubles Disability Risk in Seniors, Study Finds | 3 | 1:19 | 37 |
| 79 | https://www.youtube.com/watch?v=MxURWgdEfuY | Sedentary lifestyle p2 -It's Bad!!! This's How I solved My Problem of Chronic Sedentary Lifestyle | 28 | 4:52 | 127 |
| 80 | https://www.youtube.com/watch?v=EUG_lUEJfUI | Sedentary Lifestyle Takes Toll on Health | 3 | 2:01 | 76 |
| 81 | https://www.youtube.com/watch?v=Q36IfFpzwqY | Sedentary to Active Lifestyle | 79 | 4:55 | 883 |
| 82 | https://www.youtube.com/watch?v=I9dC2ASKT8U | Should Sedentary Lifestyle Be Considered a Medical Condition | 21 | 1:27 | 138 |
| 83 | https://www.youtube.com/watch?v=5l8w3OWC4BM | Sitting Is the New Smoking?? | 3 | 2:17 | 2577 |
| 84 | https://www.youtube.com/watch?v=Wl3U8DlGlyU | Sleekgeek Talks to Heath24 | 21 | 3:28 | 1643 |
| 85 | https://www.youtube.com/watch?v=391MFsMJeyo | Steven Needs to Change his Sedentary Lifestyle. | 52 | 3:57 | 608 |

| Result No. | URL | Title | Months on YouTube | Length (min) | Views |
|---|---|---|---|---|---|
| 86 | https://www.youtube.com/watch?v=go-tapi_c7H0 | Stop Sitting Your Life Away | 10 | 1:42 | 1294 |
| 87 | https://www.youtube.com/watch?v=2oDi1n4Cd-so | The American Sedentary Lifestyle | 40 | 1:58 | 1061 |
| 88 | http://www.youtube.com/watch?v=8h3HM9O2nQU | Thesis Defense | 8 | 1:44:46 | 235 |
| 89 | http://www.youtube.com/watch?v=-NAFN0tzjBE | To Good Health: Battling Chronic Diseases Episode 1 | 29 | 12:29 | 33 |
| 90 | http://www.youtube.com/watch?v=ysx-pX1I4VPY | Toddlers on the Move | 0 | 1:43 | 22 |
| 91 | https://www.youtube.com/watch?v=8r0tBE-TyxM | Top Health and Fitness | 20 | 1:05 | 2796 |
| 92 | http://www.youtube.com/watch?v=c2Sa1Gczhoc | Walk with Wellness | 2 | 2:02 | 25 |
| 93 | https://www.youtube.com/watch?v=FaQbr-cJU150 | Walking: Get the Word Out | 15 | 1:17 | 535 |
| 94 | http://www.youtube.com/watch?v=3lGrOyF-pjTg | We are Leading Sedentary Lives, Says Health Minister | 4 | 2:16 | 16 |
| 95 | https://www.youtube.com/watch?v=_mNe_IF8ocg | What are Risk Factors Associated with Sedentary Lifestyle and Poor Nutritional Habits in Brazil? | 35 | 1:18 | 263 |
| 96 | https://www.youtube.com/watch?v=5i5-ox64reY | What can we do to Combat a Sedentary Lifestyle? | 12 | 3:14 | 68 |
| 97 | http://www.youtube.com/watch?v=SWnOGts8Oew | Your Chair Is Killing You - Ernesto Ramirez | 39 | 5:15 | 842 |
| 98 | http://www.youtube.com/watch?v=MK9QG-GzeQe4 | Zamzee Interview on KRON 4 News | 19 | 4:45 | 301 |
| 99 | https://www.youtube.com/watch?v=iX-hEedeIRnM&list=PLbx2DInNN4q-ZXkOL23D5VQlNka4NfTwb | Physical Inactivity: The Biggest Public Health Problem of the 21st Century | 7 | 1:07:17 | 1542 |
| 100 | https://www.youtube.com/watch?v=I-RkF-DrFhF8&list=PLbx2DInNN4q-ZXkOL23D5VQlNka4NfTwb&index=2 | Too Much TV Bad for Your Health? | 35 | 1:49 | 5253 |
| 101 | https://www.youtube.com/watch?v=dEq-tySX0wSs&list=PLbx2DInNN4q8eXTUVJh-FuTYTLxhmcK-mm | Inactivity Increases Heart Disease Risk \| Heart Disease | 9 | 1:22 | 707 |
| 102 | https://www.youtube.com/watch?v=3LBbH-bZ8jxk | Report Says: Physical Inactivity Kills 5 Million a Year | 22 | 7:19 | 108 |
| 103 | https://www.youtube.com/watch?v=zrp5sF0za40 | Dr Rutledge Cause of Obesity: Excessive Calories/Lack of Exercise. Calories In and Calorie Out | 51 | 5:51 | 720 |
| 104 | https://www.youtube.com/watch?v=I9dC2ASKT8U | Should Sedentary Lifestyle Be Considered a Medical Condition | 21 | 1:27 | 361 |
| 105 | https://www.youtube.com/watch?v=kPN-MuuY7jLU | The Relationship Between Physical Activity and Childhood Overweight and Obesity | 42 | 1:55 | 84 |
| 106 | https://www.youtube.com/watch?v=Ofg3UlxFVM0 | Couch Potato Toddlers | 38 | 1:09 | 635 |

**Table 2.** Location of origin for videos (N=106).

| Country | Videos, n (%) |
|---|---|
| Australia | 4 (3.7) |
| Barbados | 1 (0.9) |
| Belgium | 1 (0.9) |
| Canada | 12 (11.3) |
| Colombia | 1 (0.9) |
| Kenya | 1 (0.9) |
| New Zealand | 1 (0.9) |
| Russia | 1 (0.9) |
| South Africa | 1 (0.9) |
| Spain | 2 (1.9) |
| Ukraine | 1 (0.9) |
| United Kingdom | 8 (7.6) |
| United States | 44 (41.5) |
| Not available | 28 (26.4) |

**Figure 1.** Search results.

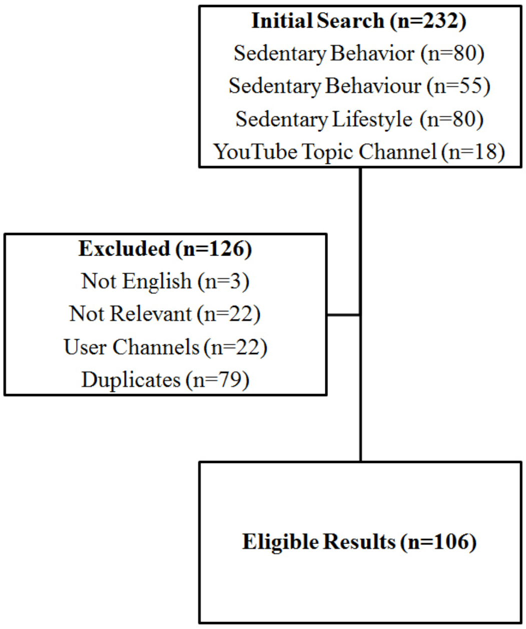
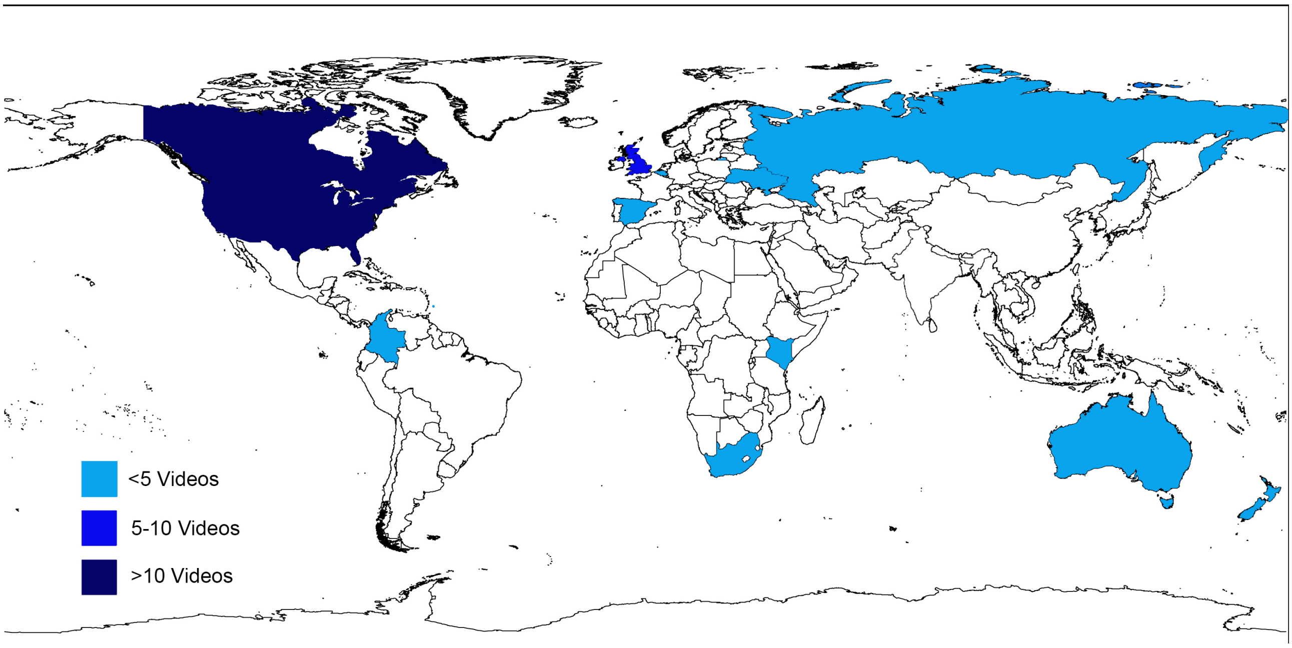


**Figure 2.** Global distribution of search results.



## Video Features

The four search phrases ("sedentary behavior", "sedentary behaviour", "sedentary lifestyle", and YouTube topic channel) and two search methods (sorted by relevance or views) generated 58.5% (62/106) duplicate results. Of the results that were unique to a single search phrase (41.5%, 44/106), "sedentary lifestyle" generated 65.9% (29/44), "sedentary behaviour" generated 20.5% (9/44), and "sedentary behavior" generated 13.6% (6/44).

Table 3 presents the means and interquartile ranges of video views, length, time since uploaded to the YouTube site, and quantity of comments posted by users. There was no relationship between number of views and video length ($r$=-.10, $P$>.05), or the search phrase ($r$=.08, $P$>.05). Similarly, no relationship was evident between search phrase and the source of the video ($r$=-.05, $P$>.05) or the activity type presented in the video ($r$=-.09, $P$>.05).

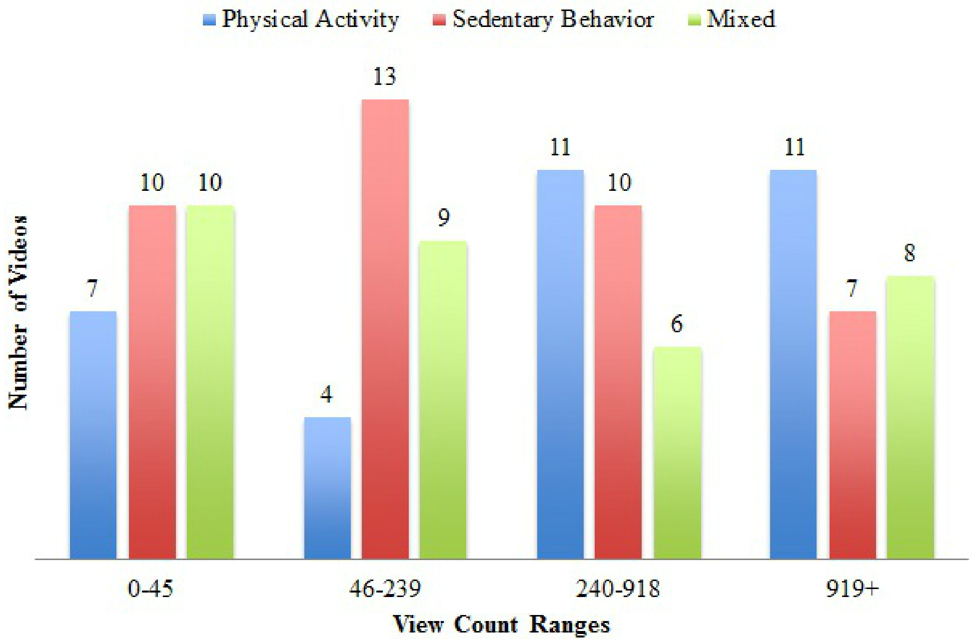
**Table 3.** Descriptive features of video results (N=106).

| | Mean (SD) | 25th percentile | 50th percentile | 75th percentile |
|---|---|---|---|---|
| Views | 1008.9 (2202) | 44.5 | 239.0 | 917.5 |
| Video length (minutes) | 9:01 (16:31) | 1:44 | 3:00 | 5:40 |
| Time on YouTube (months) | 19.2 (17.0) | 6.0 | 15.0 | 27.5 |
| Comments | 1.7 (5.7) | 0 | 1 | 1 |
| Shares | 1.2 (3.9) | 0 | 0 | 1 |

## Activity Type

Videos that portrayed content on sedentary behaviors alone represented 37.7% (40/106) of the results. Videos that portrayed a combination of sedentary and physical activity behaviors represented 31.1% (33/106) of the sample, and videos that portrayed physical activity behaviors alone similarly comprised 31.1% (33/106) of the sample. Videos with the highest view counts (ie >240 views) portrayed physical activity behavior content, not specifically sedentary behavior content (Figure 3).

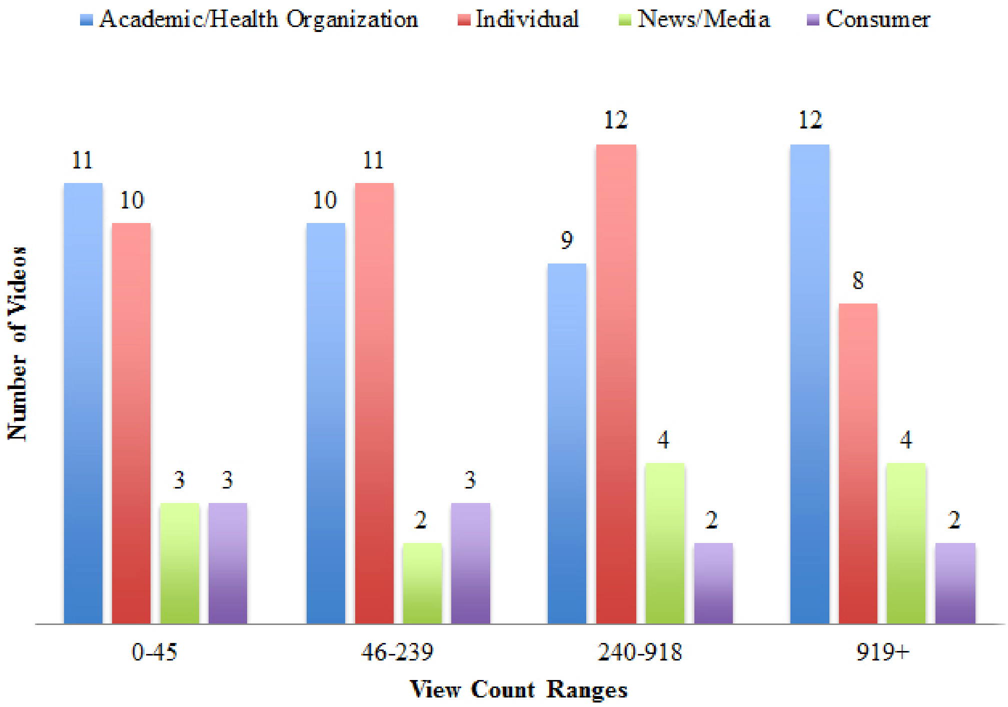
**Figure 3.** Activity type portrayed in videos by interquartile range of views.



## Source of the Content

Academic/Health Organizations and Individuals were the most common source of content, representing 39.6% (42/106) and 38.7% (41/106) of videos, respectively. News/Media comprised 12.3% (13/106) of the videos, and 9.4% (10/106) of videos were Consumer. The trend was similar across view count ranges (Figure 4).

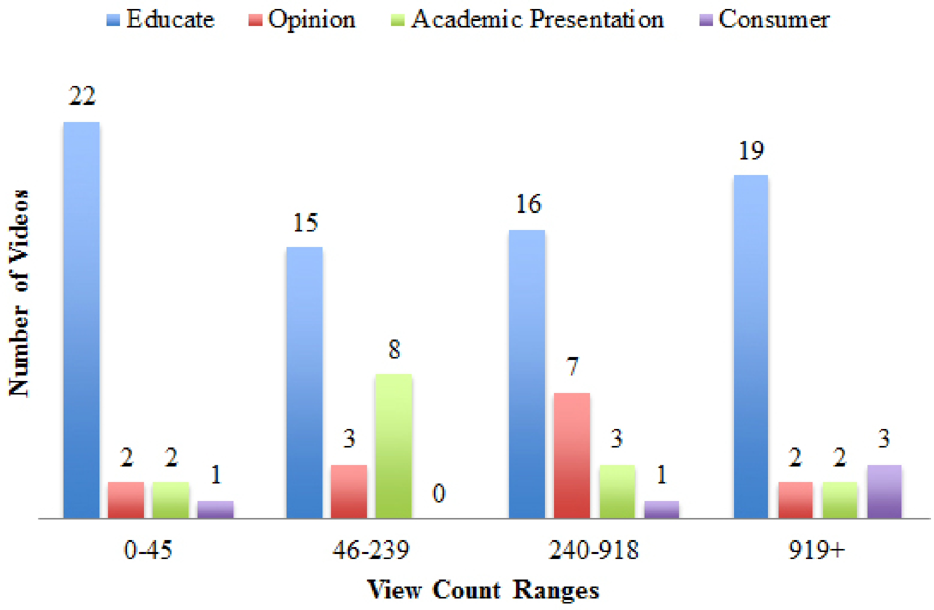
**Figure 4.** Source of videos presented by interquartile range of views.



## Purpose of the Content

The predominant purpose of videos was coded as Educational (71/106, 67.0%). Academic presentations (16/106, 15.1%) and Opinions (14/106, 13.2%) contributed to a smaller portion of the available content. Minimal videos (5/106, 4.7%) were Consumer-based, aiming to sell products or services. Moreover, Educational videos were more dominant across all view count ranges (Figure 5).

**Figure 5.** Purpose of video presented by interquartile range of views.



## Discussion

### Popularity of Content

YouTube is a popular social media site that facilitates the sharing of content, evidence-based or otherwise, with a large body of users. The purpose of this study was to explore descriptive features of sedentary behavior content on YouTube to enhance future knowledge translation efforts of evidence-informed sedentary behavior content. Results from 106 sedentary behavior videos demonstrate that content is being uploaded from around the world, which further underscores the value of this medium as a way to link the global community. Moreover, these results may underestimate the global reach of content via this social platform due to eligibility criteria from this study excluding videos not available in English.

Google AdWords [27] was used to generate a list of search phrases common to Internet users. The search phrase “sedentary lifestyle” generated more unique search results on YouTube than the other search phrases. However, no relationship was evident between search phrase and either views or source of the video. These results suggest that certain key terms are more prevalent on the site, yet this does not appear to impact viewership of sedentary behavior content on this social media platform.

To help design the future development of sedentary behavior content for YouTube, our investigation explored the relationship between length of videos and view counts. There was no relationship, suggesting that length of video does not impact the viewership. Therefore, knowledge translation efforts may need to consider aspects other than length of videos in order to increase viewership.

User-generated comments on videos represent the interactive and collaborative nature of YouTube content. The videos included this analysis had generated very few comments on the YouTube site (Table 3), which may indicate that sedentary behavior content currently on YouTube is not generating discussion and collaboration among users. Similarly, there were a remarkably low number of shares for sedentary behavior videos, indicating that users are not engaging in further social features of the YouTube site to enhance the reach of sedentary behavior content.

Moreover, view counts can be used to compare popularity of content on YouTube. Typically, popular videos on the site generate hundreds of thousands to millions of views. The median view count of sedentary behavior videos was 239 (Table 3), which suggests that the sedentary behavior content posted to the site is not popular among users. This presents a substantial opportunity for sedentary behavior researchers to improve the reach and impact of evidence through this social platform.

### Evidence-Based Content

In the past, the term “sedentary” was often used to refer to individuals who were not sufficiently physically active [4,29]. However, as noted above, available evidence suggests that sedentary behavior and physical activity should be viewed as separate and distinct constructs [4,30]. Thus, a growing number of researchers have suggested that the term “sedentary” should be used only to describe sedentary behaviors (eg, activities done while sitting), as opposed to the lack of physical activity [4].

Despite the widely used academic definition of sedentary behavior as activities characterized by sitting, approximately one third of this sample of videos displayed content on physical activity, not sedentary behaviors. Further, videos with higher view counts tended to portray information on physical activity, not sedentary behaviors (Figure 3). Finally, there was no relationship between search phrase and the type of activity behavior presented in the video (ie, sedentary behavior, physical activity, or mixed). These results further underscore the confusion between physical activity and sedentary behavior, which may impact knowledge users’ understanding of both the behaviors themselves and the associated health outcomes of these distinct behaviors.

### Limitations

The source and content of information is variable in this medium. There is currently a lack of standardized tools for assessing quality of content on social media sites like YouTube. Unlike systematic reviews of traditional evidence, a gap exists in the literature describing an evidence-based quality assessment tool for the purpose of reviewing social media content.

Moreover, previous research has highlighted concerns about the regulation of content available online. For example, one study that examined YouTube for evidence-based immunization content found that videos containing information that contradicted public health guidelines on the topic of interest were more likely to receive high view counts and user ratings and accounted for more than half of YouTube content on the topic [20]. Moreover, another study that examined YouTube for video content on eating disorders found that one third of videos glorified the unhealthy behaviors, and that these videos were more likely to have higher view counts than videos that discouraged the behavior [25]. These findings may further underscore the importance of the research community leveraging the popularity of YouTube as a knowledge translation vehicle to promote evidence-based information.

While YouTube is accessible around the world, and results from this study demonstrate content being uploaded from across the globe, many workplaces and educational institutions restrict user access to social media sites like YouTube. Restricted access may limit the impact of this medium for translation of evidence-based information to users. Research demonstrating the cultural value of YouTube [31,32] in conjunction with the potential of the social platform for sharing evidence-based content may be of value for informing the future development of policies governing access to social sites like YouTube.

### Conclusions

While physical activity is a well-established research domain, the focus on sedentary behavior research is much more recent. Therefore, many information seekers are not only unaware of the health consequences of prolonged sedentary behaviors, but also of the distinction between “too much sitting” and “not enough exercise”. Moreover, there is a shortage of evidence describing the implementation and translation of evidence-based sedentary behavior information into lay landscapes, which could further impede an individual’s understanding of this health risk.

Our study explored sedentary behavior content available on YouTube. Findings demonstrate that there is confusion between physical activity and sedentary behaviors, that content is being uploaded to the site from around the globe, that content is primarily from health organizations and individuals with the purpose of educating fellow users, but that low views and comments suggest that sedentary behavior content is relatively underutilized on YouTube. Future research may wish to leverage social platforms, such as YouTube, to facilitate implementation of evidence-based sedentary behavior content.

## Acknowledgments

EK is supported by a Canadian Institutes of Health Research Doctoral Research Award. Sponsors had no role in the development of this project or the submission of this manuscript for publication.

## Authors' Contributions

EK contributed to research design, implementation, data collection, data analysis, and manuscript preparation. BI contributed to data collection, analysis, and manuscript preparation. AM contributed to data collection, analysis, and manuscript preparation. TJS contributed to research design, data analysis, and manuscript preparation. All authors approved the manuscript for publication.

## Conflicts of Interest

TJS was the creator of one of the videos included in this analysis.

## References

1. Morris J, Crawford M. Coronary Heart Disease and Physical Activity of Work. BMJ 1958 Dec 20;2(5111):1485-1496. [doi: 10.1136/bmj.2.5111.1485]
2. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
3. Global Health Recommendations on Physical Activity for Health. Geneva: World Health Organization; 2010.
4. Sedentary Behaviour Research Network. Letter to the Editor: Standardized use of the terms "sedentary" and "sedentary behaviours". Appl Physiol Nutr Metab 2012 Jun;37(3):540-542. [doi: 10.1139/H2012-024]
5. Hamilton MT, Hamilton DG, Zderic TW. Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. Diabetes 2007 Nov;56(11):2655-2667 [FREE Full text] [doi: 10.2337/db07-0882] [Medline: 17827399]
6. Owen N, Healy GN, Matthews CE, Dunstan DW. Too much sitting: the population health science of sedentary behavior. Exerc Sport Sci Rev 2010 Jul;38(3):105-113 [FREE Full text] [doi: 10.1097/JES.0b013e3181e373a2] [Medline: 20577058]
7. Tremblay MS, Colley RC, Saunders TJ, Healy GN, Owen N. Physiological and health implications of a sedentary lifestyle. Appl Physiol Nutr Metab 2010 Dec;35(6):725-740. [doi: 10.1139/H10-079] [Medline: 21164543]

8. Saunders TJ, Larouche R, Colley RC, Tremblay MS. Acute sedentary behaviour and markers of cardiometabolic risk: a systematic review of intervention studies. J Nutr Metab 2012;2012:712435 [FREE Full text] [doi: 10.1155/2012/712435] [Medline: 22754695]
9. Dunstan DW, Kingwell BA, Larsen R, Healy GN, Cerin E, Hamilton MT, et al. Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care 2012 May;35(5):976-983 [FREE Full text] [doi: 10.2337/dc11-1931] [Medline: 22374636]
10. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, et al. Breaks in sedentary time: beneficial associations with metabolic risk. Diabetes Care 2008 Apr;31(4):661-666. [doi: 10.2337/dc07-2046] [Medline: 18252901]
11. Dunstan DW, Healy GN, Sugiuama T, Owen N. Too much sitting' and metabolic risk - Has modern technology caught up with us? European Endocrinology 2010;6(1):19-23.
12. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical Activity of Canadian Adults: Accelerometer Results from the 2007 to 2009 Canadian Health Measures Survey. Health Reports 2011;22(1):1-8.
13. Katzmarzyk PT, Lee IM. Sedentary behaviour and life expectancy in the USA: a cause-deleted life table analysis. BMJ Open 2012;2(4) [FREE Full text] [doi: 10.1136/bmjopen-2012-000828] [Medline: 22777603]
14. Veerman JL, Healy GN, Cobiac LJ, Vos T, Winkler EA, Owen N, et al. Television viewing time and reduced life expectancy: a life table analysis. Br J Sports Med 2012 Oct;46(13):927-930. [doi: 10.1136/bjsports-2011-085662] [Medline: 23007179]
15. Graham ID, Logan J, Harrison MB, Straus SE, Tetroe J, Caswell W, et al. Lost in knowledge translation: time for a map? J Contin Educ Health Prof 2006;26(1):13-24. [doi: 10.1002/chp.47] [Medline: 16557505]
16. Fisher J, Clayton M. Who gives a tweet: assessing patients' interest in the use of social media for health care. Worldviews Evid Based Nurs 2012 Apr;9(2):100-108. [doi: 10.1111/j.1741-6787.2012.00243.x] [Medline: 22432730]
17. Gold J, Pedrana AE, Stoove MA, Chang S, Howard S, Asselin J, et al. Developing health promotion interventions on social networking sites: recommendations from The FaceSpace Project. J Med Internet Res 2012;14(1):e30 [FREE Full text] [doi: 10.2196/jmir.1875] [Medline: 22374589]
18. Thackeray R, Crookston BT, West JH. Correlates of health-related social media use among adults. J Med Internet Res 2013;15(1):e21 [FREE Full text] [doi: 10.2196/jmir.2297] [Medline: 23367505]
19. Heilman JM, Kemmann E, Bonert M, Chatterjee A, Ragar B, Beards GM, 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]
20. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
21. Alexa: The Web Information Company. 2013. URL: http://www.alexa.com/siteinfo/youtube.com [accessed 2014-08-20] [WebCite Cache ID 6RxNQ5EaE]
22. Azer SA, Algrain HA, AlKhelaif RA, AlEshaiwi SM. Evaluation of the educational value of YouTube videos about physical examination of the cardiovascular and respiratory systems. J Med Internet Res 2013;15(11):e241 [FREE Full text] [doi: 10.2196/jmir.2728] [Medline: 24225171]
23. Knight E. Leveraging the Power of Social Media to Engage Patients and Achieve Treatment Outcomes. Crit Rev Phys Rehabil Med 2012;24(3-4):169-177. [doi: 10.1615/CritRevPhysRehabilMed.2013007169]
24. Health Care Social Media List. URL: http://network.socialmedia.mayoclinic.org/hcsml-grid [WebCite Cache ID 6RxPSI6h7]
25. Syed-Abdul S, Fernandez-Luque L, Jian WS, Li YC, Crain S, Hsu MH, et al. Misleading health-related information promoted through video-based social media: anorexia on YouTube. J Med Internet Res 2013;15(2):e30 [FREE Full text] [doi: 10.2196/jmir.2237] [Medline: 23406655]
26. Williams D, Sullivan SJ, Schneiders AG, Ahmed OH, Lee H, Balasundaram AP, et al. Big hits on the small screen: an evaluation of concussion-related videos on YouTube. Br J Sports Med 2014 Jan;48(2):107-111. [doi: 10.1136/bjsports-2012-091853] [Medline: 23446643]
27. Google Adwords Keyword Planner. URL: https://adwords.google.com/KeywordPlanner [accessed 2014-08-20] [WebCite Cache ID 6RxQHR0b9]
28. YouTube Stats (Subscriber Ranking & Video Statistics). URL: http://vidstatsx.com/ [accessed 2014-08-21] [WebCite Cache ID 6RyvXIyCa]
29. Pate RR, O'Neill JR, Lobelo F. The evolving definition of "sedentary". Exerc Sport Sci Rev 2008 Oct;36(4):173-178. [doi: 10.1097/JES.0b013e3181877d1a] [Medline: 18815485]
30. Saunders TJ, Chaput JP, Tremblay MS. Sedentary behaviour as an emerging risk factor for cardiometabolic diseases in children and youth. Can J Diabetes 2014 Feb;38(1):53-61. [doi: 10.1016/j.jcjd.2013.08.266] [Medline: 24485214]
31. Wesh M. Web 2.0...The machine is us/ing us. URL: https://www.youtube.com/watch?v=6gmP4nk0EOE [accessed 2014-08-20] [WebCite Cache ID 6RxMWAK8Q]
32. Wesh M. An anthropological introduction to YouTube. 2008. URL: https://www.youtube.com/watch?v=TPAO-lZ4_hU [accessed 2014-08-20] [WebCite Cache ID 6RxLuyCbd]

*Edited by G Eysenbach; submitted 02.09.14; peer-reviewed by S Sullivan, DM Liou; comments to author 22.09.14; revised version received 02.10.14; accepted 04.11.14; published 20.01.15.*

*Please cite as:*
*Knight E, Intzandt B, MacDougall A, Saunders TJ*
*Information Seeking in Social Media: A Review of YouTube for Sedentary Behavior Content*
*Interact J Med Res 2015;4(1):e3*
*URL: http://www.i-jmr.org/2015/1/e3/*
*doi:10.2196/ijmr.3835*
*PMID:25604433*

Original Paper

# Content and Quality of Information Websites About Congenital Heart Defects Following a Prenatal Diagnosis

Tommy Carlsson[1], RN, PGDip, MMSc; Gunnar Bergman[2], MD, PhD; Anna-Malin Karlsson[3], PhD; Elisabet Mattsson[1,4], RNM, PhD

[1]Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden
[2]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden
[3]Department of Scandinavian Languages, Uppsala University, Uppsala, Sweden
[4]Department of Women's and Children's Health, Uppsala University, Uppsala, Sweden

**Corresponding Author:**
Tommy Carlsson, RN, PGDip, MMSc
Department of Public Health and Caring Sciences
Uppsala University
BMC Husargatan 3
Box 564
Uppsala, S-75122
Sweden
Phone: 46 736143454
Fax: 46 184716675
Email: tommy.carlsson@pubcare.uu.se

## Abstract

**Background:** Pregnant women and their partners use the Internet to search for information following a prenatal diagnosis of congenital heart defect.

**Objective:** Our aim was to explore central subjects of content and to assess the accessibility, reliability, usability, and quality of written information on publicly available information websites about congenital heart defects following a prenatal diagnosis.

**Methods:** Following searches on Bing and Google, we included websites containing patient information in English. Hits ranged from 340,000-67,500,000 and the first 50 hits from each search were screened for inclusion (N=600). Of these hits, 39.3% (236/600) were irrelevant. A total of 67 websites were included, of which 37% (25/67) were affiliated with independent information websites, 25% (17/67) with charity/private organizations, 25% (17/67) with hospitals/clinics, and 13% (8/67) had other affiliations. The majority of the websites (76%, 51/67) could not be attributed to an author. A manifest content analysis was performed to explore central subjects of content. The DISCERN instrument was used to assess the quality of information, and the LIDA tool was used to assess accessibility, usability, and reliability of the included websites.

**Results:** The content on the majority of the websites included care and treatment of children with congenital heart defects (88%, 59/67), causes of congenital heart defects (88%, 59/67), symptoms of congenital heart defects (85%, 57/67), prevalence of congenital heart defects (81%, 54/67), potential complications of congenital heart defects (75%, 50/67), prenatal diagnostics/screening methods (72%, 48/67), and specific congenital heart defects (72%, 48/67), whereas less than 10% included information about termination of pregnancy (6%, 4/67), care during pregnancy (5%, 3/67), and information specifically directed to partners (1%, 1/67). The mean of the total DISCERN score was 27.9 (SD 9.7, range 16-53). According to the instrument, a majority of the websites were categorized as very poor regarding information about effects of no treatment (88%, 59/67), support for shared decision making (85%, 57/67), achievement of its aims (84%, 56/67), explicit aims (82%, 55/67), risks of each treatment (82%, 55/67), how treatment choices affect overall quality of life (76%, 51/67), and areas of uncertainty (76%, 51/67). The mean of the total LIDA score was 92.3 (SD 13.1, range 61-127). According to the tool, a majority of the websites were categorized as good with regard to registration (97%, 65/67) and browser test (75%, 50/67), whereas a majority were categorized as poor with regard to currency (87%, 58/67), content production (84%, 56/67), and engagability (75%, 50/67).

**Conclusions:** Difficulties in finding relevant information sources using Web search engines and quality deficits on websites are an incentive for health professionals to take an active part in providing adequate and reliable information online about congenital heart defects.

***(Interact J Med Res 2015;4(1):e4)*** doi:10.2196/ijmr.3819

**KEYWORDS**

consumer health information; heart defects, congenital; Internet; prenatal diagnosis

## Introduction

Globally, an increasing number of health care consumers use the Internet to search for health-related issues [1-4]. The Internet has the potential to provide highly accessible, interactive, and tailored information. However, this might be limited by navigational difficulties and inaccurate or misleading information that has not been peer-reviewed [5,6]. Although many individuals have little or no trust in Internet information, it is used as a primary source when searching for health-related information [7]. Studies of literacy practices in relation to health communication have shown that trustworthiness is the key issue for patients assessing health information. For example, pregnant women in the United Kingdom were shown to search for texts written by medical professionals or published by medical institutions, thus appraising authority on the basis of their trust in academic and professional expertise [4].

Advances in prenatal screening have improved the detection rate of fetal diagnoses of congenital heart defects (CHD) [8,9]. Following diagnosis, counseling from health professionals is essential regarding a wide range of topics, including, for example, the nature and consequences of the CHD, severity, treatments available, prognosis, postoperative complications, and possible associations of CHD with other diseases [10,11]. Based on the information received, the pregnant woman also has the option of and the responsibility for deciding whether or not to terminate the pregnancy. Depending on national legislation on termination of the pregnancy, the decision must often be made soon after receiving the diagnosis. The process towards an informed decision on the future of the pregnancy involves various difficulties, including comprehending complex medical information [12,13], ethical considerations [14,15] and psychological distress [16]. An online survey among parents of children with CHD revealed that 50% report that more information at the time of diagnosis would have been helpful [17]. To deal with this matter, pregnant women and their partners try to supplement counseling from health professionals by using the Internet to search for information following the diagnosis [12].

The aim of this study was to explore central subjects of content and to assess the accessibility, usability, reliability, and quality of written information on publicly available information websites about congenital heart defects following a prenatal diagnosis.

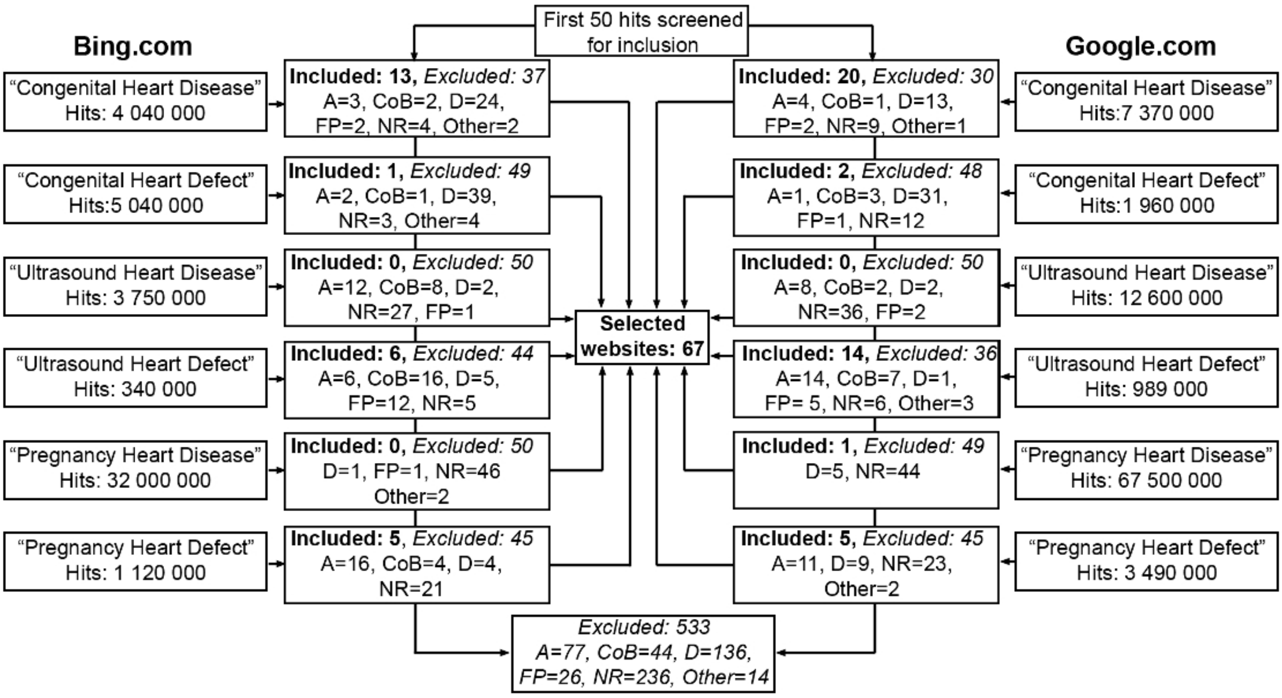
## Methods

### Data Collection

In October 2013, the following key terms were entered separately in the two most commonly used search engines, Bing and Google [18]: “Congenital Heart Disease”, “Congenital Heart Defect”, “Ultrasound Heart Disease”, “Ultrasound Heart Defect”, “Pregnancy Heart Disease”, and “Pregnancy Heart Defect”. The inclusion criterion was a website written in English that provided patient information regarding CHD. The search was made in incognito mode in order to minimize influence from previous search patterns. The generated result of the search procedure was saved, and the first 50 hits obtained for each search procedure were screened for inclusion (N=600). Duplicate websites and direct links to communities/blogs, video materials, and scientific articles were excluded. In total, 533 (88.8%) were excluded, leaving 67 (11.2%) websites for inclusion in the study. Figure 1 presents the selection process, that is, key terms, hits, and excluded and included websites via Bing and Google, respectively.

The websites we included were affiliated with independent information websites (25/67, 37%), charity/private organizations (17/67, 25%), hospitals/clinics (17/67, 25%), governments (4/67, 6%), medical companies (1/67, 2%), and other websites (4/67, 6%). The majority of the websites could not be attributed to an author (51/67, 76%), whereas a minority could be attributed to medical professionals (10/67, 15%), journalists (2/67, 3%), and others (4/67, 6%).

**Figure 1.** Selection process for the included websites (N=67; A: Scientific Article; CoB: Community or Blog; D: Duplicate; FP: For Professionals; NR: Not Relevant).



## Data Analyses

### Central Subjects of Content

A manifest content analysis [19] was performed to inductively identify the central subjects of content of each website. The first author read the websites repeatedly. Paragraphs and statements (meaning units) containing relevant information regarding the content of each website were identified and grouped into categories (central subjects). Meaning units in the same category are assumed to have a similar meaning, on the basis of either the precise meaning of text or of texts sharing the same connotations. Thereafter, the websites were read again and subcategories were identified, that is, common characteristics within a larger category. Finally, the websites were all read through once more in order to validate the results.

### Assessment of Quality

The websites were individually evaluated using two standardized instruments: the DISCERN instrument [20] and the Minervation validation instrument for health care websites (LIDA tool, version 1.2) [21]. The first author conducted all assessments of the websites.

The DISCERN instrument is a reliable and valid instrument for assessing the quality of written consumer health information independent of previous knowledge of the field under research [20]. It was developed with the input of an expert panel, health information providers, and patients from a self-help group, and has acceptable levels of interrater agreement [22]. The instrument consists of 16 questions divided into three sections. The user rates each question on a scale ranging from 1 (low/poor) to 5 (high/excellent), resulting in a total score ranging from 16 to 80. Section 1 includes 8 questions (score ranging from 8-40) and assesses reliability, whereas Section 2 with 7 questions (7-35) focuses on the quality of information about treatment options, that is, in this study continuation/termination of pregnancy, fetal interventions, and treatments of CHD. Section 3 consists of 1 question (1-5) and provides an overall rating of the quality of the websites, based on the responses to the previous questions [20].

The LIDA tool [21] assesses accessibility, reliability, and usability of health care websites. It consists of 29 questions and an automated test. Each question is rated from 0 to 3 (higher being better), and the automated test generates a score of 0-54, resulting in a total score ranging from 0 to 141. The overall score is calculated as a percentage, where scores greater than 90% represent good results and less than 50% poor results [23]. Accessibility (score ranging from 0-60) includes an automated test of page set-up, access restrictions, and outdated code, together with manual registration and browser tests conducted in Apple Safari, Google Chrome, and Mozilla Firefox. Usability (0-54) includes clarity, consistency, functionality, and engagability. Reliability (0-27) includes currency, conflicts of interest, and content production.

### Statistical Analysis

Descriptive statistics were carried out using R (version 3.0.1).

# Results

## Central Subjects of Content

Central subjects of content on the 67 websites were categorized into 25 categories with 46 subcategories (Table 1).

**Table 1.** Central subjects of content identified on the included websites (N=67).

| Category | Subcategory | n (%) |
|---|---|---|
| **Care and treatment of children with CHD** | | 59 (88) |
| | Surgery | 59 (88) |
| | Cardiac catheterization for treatment | 48 (72) |
| | Medications | 48 (72) |
| | Cardiac transplantation | 39 (58) |
| | Nutrition | 18 (27) |
| | Pacemaker | 17 (25) |
| | Intensive care | 11 (16) |
| | Animations or illustrations of treatments | 9 (13) |
| | Immunizations | 3 (5) |
| | How to include cultural/spiritual beliefs in the care of the child | 1 (1) |
| Causes of CHD | | 59 (88) |
| Symptoms of CHD | | 57 (85) |
| Prevalence of CHD | | 54 (81) |
| Potential complications of CHD | | 50 (75) |
| **Prenatal diagnostic/screening methods** | | 48 (72) |
| | Fetal echocardiography | 40 (60) |
| | Amnioscentesis | 15 (22) |
| | Chorionic villus sampling | 12 (18) |
| | Nuchal translucency scan | 11 (16) |
| | Blood tests | 9 (13) |
| | Risks of invasive methods | 9 (13) |
| | Umbilical cord sampling | 4 (6) |
| | Fetal magnetic resonance imaging | 2 (3) |
| **Specific CHD** | | 48 (72) |
| | Animations or illustrations of CHD | 28 (42) |
| Associated anomalies | | 47 (70) |
| **Normal cardiovascular system** | | 46 (69) |
| | Postnatal cardiovascular system | 42 (63) |
| | Cardiovascular changes at birth | 35 (52) |
| | Animations or illustrations of normal cardiovascular system | 29 (43) |
| | Fetal cardiovascular system | 16 (24) |
| **Postnatal diagnostic methods** | | 45 (67) |
| | Echocardiography | 43 (64) |
| | Electrocardiography | 42 (63) |
| | Physical examination | 42 (63) |
| | Radiography | 42 (63) |
| | Cardiac catheterization for diagnosis | 39 (58) |
| | Pulse oximeter | 25 (37) |
| | Chemical analyses | 18 (27) |
| | Exercise test | 16 (24) |
| **Long-term outlook and care** | | 44 (66) |

| Category | Subcategory | n (%) |
|---|---|---|
| | Monitoring/Follow-up care | 31 (46) |
| | Dental care/endocarditis prophylaxis | 28 (42) |
| | Grown-up with CHD | 24 (36) |
| | Pregnancy with CHD in mother | 24 (36) |
| | Physical activity | 23 (34) |
| Prognosis | | 41 (61) |
| Risks of treatment of CHD | | 25 (37) |
| Common feelings following prenatal diagnosis of CHD | | 18 (27) |
| **Postnatal quality of life** | | 18 (27) |
| | Quality of life for the child | 18 (27) |
| | Quality of life for the family | 5 (7) |
| Examples of previous cases that continued the pregnancy | | 16 (24) |
| Precision of prenatal diagnosis of CHD | | 16 (24) |
| **Delivery** | | 14 (21) |
| | Location and planning of delivery | 14 (21) |
| | Mode of delivery | 1 (1) |
| **Postnatal coping with the diagnosis** | | 13 (19) |
| | Financial issues | 12 (15) |
| | Grief and bereavement | 2 (3) |
| | Information regarding siblings | 2 (3) |
| Risks of CHD in future pregnancy | | 10 (15) |
| Fetal intervention | | 8 (12) |
| Presentation of the multidisciplinary team in care of the child | | 8 (12) |
| **Termination of pregnancy** | | 4 (6) |
| | Informed and personal decision | 2 (3) |
| | Time limit | 2 (3) |
| | Feelings about termination of pregnancy | 1 (1) |
| Care during pregnancy | | 3 (5) |
| Information specifically directed to partners | | 1 (1) |

The majority (>70%) of the websites contained information about care and treatment of children with CHD (88%, 59/67), causes of CHD (88%, 59/67), symptoms of CHD (85%, 57/67), prevalence of CHD (81%, 54/67), potential complications of CHD (75%, 50/67), prenatal diagnostics/screening methods (72%, 48/67), and specific CHD (72%, 48/67). A minority (<30%) of the websites contained information about common feelings following prenatal diagnosis of CHD (27%, 18/67), postnatal quality of life (27%, 18/67), examples of previous cases that continued the pregnancy (24%, 16/67), precision of prenatal diagnosis of CHD (24%, 16/67), delivery (21%, 14/67), postnatal coping with the diagnosis (19%, 13/67), risks of CHD in future pregnancy (15%, 10/67), fetal intervention (12%, 8/67), presentation of the multidisciplinary team in care of the child (12%, 8/67), termination of pregnancy (6%, 4/67), care during pregnancy (5%, 3/67), and information specifically directed to partners (1%, 1/67).

## Assessment of Quality

### *DISCERN*

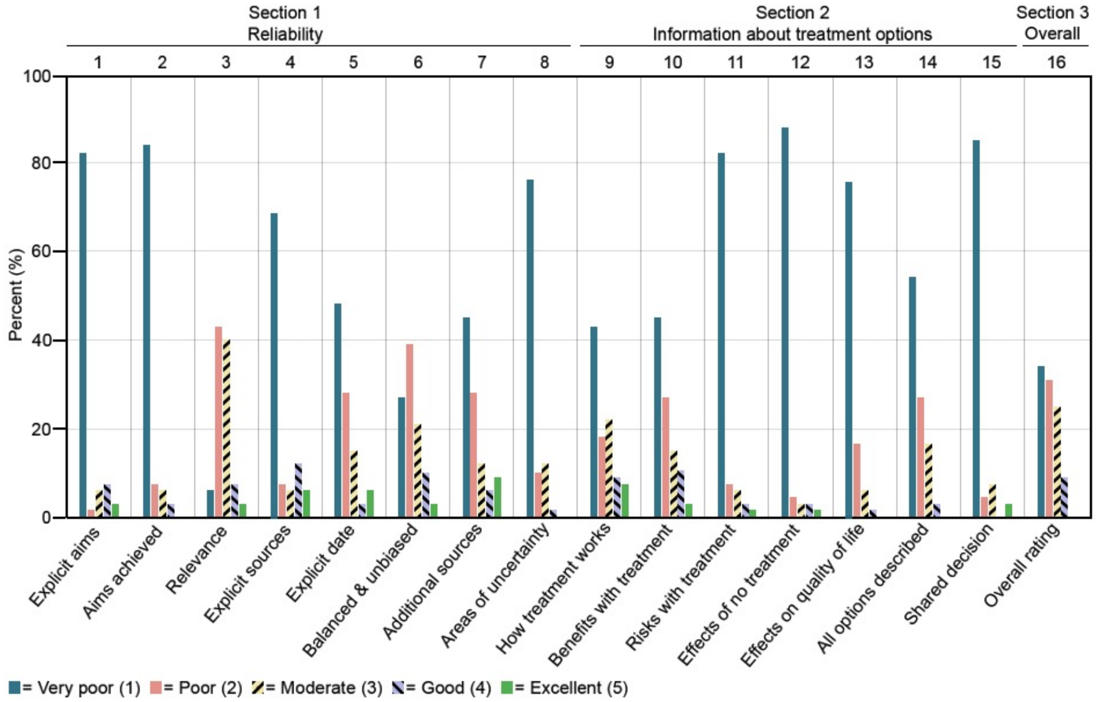
Table 2 presents means, standard deviations, and ranges of the included websites measured by the DISCERN instrument.

In applying the DISCERN criteria to the evaluation of the websites (Figure 2), the majority of the websites (>70%) were categorized as very poor regarding effects of no treatment (88%, 59/67), support for shared decision making (85%, 57/67), achievement of its aims (84%, 56/67), explicit aims (82%, 55/67), risks of each treatment (82%, 55/67), how treatment choices affect overall quality of life (76%, 51/67), and areas of uncertainty (76%, 51/67).

**Table 2.** Means, standard deviations (SD), and ranges of the included websites (N=67) measured by the DISCERN instrument (the maximum achievable scores in shown in brackets after each section and question).

| Section (max. score) | Question | Mean (SD) | Range |
|---|---|---|---|
| **Reliability (40)** | | 14.7 (5.2) | 8-30 |
| | Explicit aims (5) | 1.5 (1.1) | 1-5 |
| | Aims achieved (5) | 1.3 (0.7) | 1-4 |
| | Relevance (5) | 2.6 (0.8) | 1-5 |
| | Explicit sources (5) | 1.8 (1.3) | 1-5 |
| | Explicit date (5) | 1.9 (1.1) | 1-5 |
| | Balanced and unbiased (5) | 2.2 (1.1) | 1-5 |
| | Additional sources(5) | 2.1 (1.3) | 1-5 |
| | Areas of uncertainty (5) | 1.4 (0.8) | 1-4 |
| **Treatment options (35)** | | 11.1 (4.9) | 7-25 |
| | How treatment works (5) | 2.2 (1.3) | 1-5 |
| | Benefits of treatment (5) | 2.0 (1.1) | 1-5 |
| | Risks of treatment (5) | 1.3 (0.8) | 1-5 |
| | Effects of no treatment (5) | 1.3 (0.8) | 1-5 |
| | Effects on quality of life (5) | 1.3 (0.7) | 1-4 |
| | All options described (5) | 1.7 (0.9) | 1-4 |
| | Shared decision (5) | 1.3 (0.9) | 1-5 |
| Overall rating (5) | | 2.1 (1.0) | 1-4 |
| Total (80) | | 27.9 (9.7) | 16-53 |

**Figure 2.** Proportion of websites (N=67) categorized into five categories from very poor to excellent, for each separate question in the DISCERN instrument.



### *LIDA*

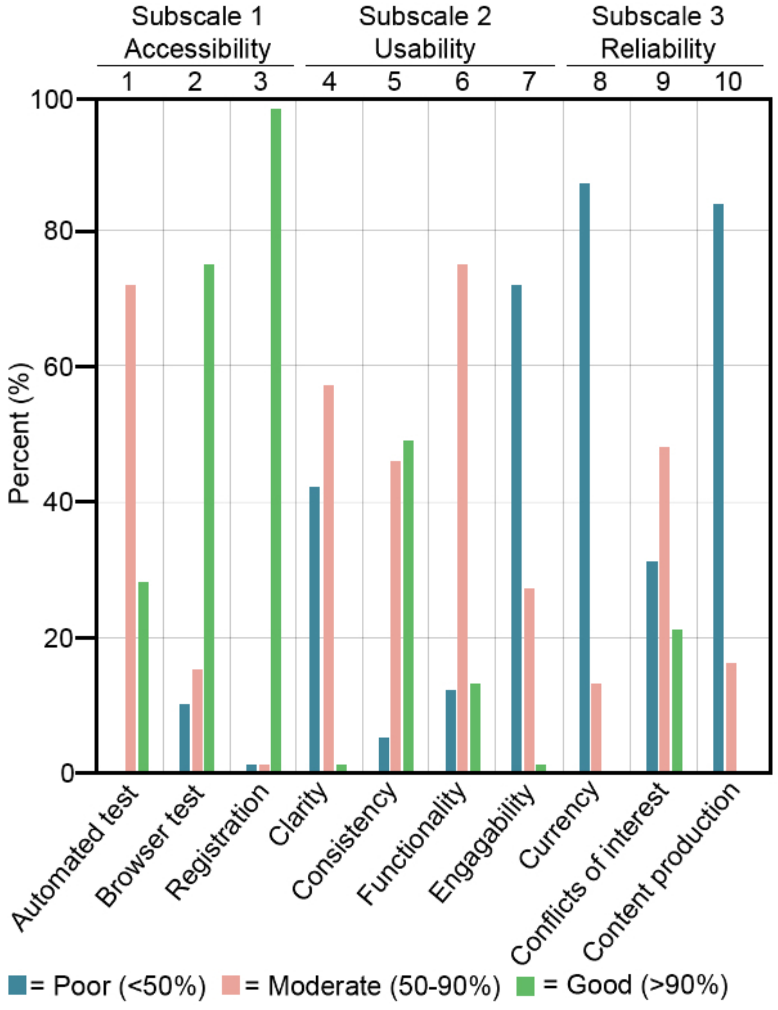
[Table 3](#) presents means, standard deviations (SD), and ranges of the included websites measured by the LIDA tool.

In applying the LIDA criteria to the evaluation of the websites ([Figure 3](#)), the majority of the websites (>70%) were categorized as good regarding registration (97%, 65/67) and browser test (75%, 50/67), whereas the majority was categorized as poor regarding currency (87%, 58/67), content production (84%, 56/67), and engagability (75%, 50/67).

**Table 3.** Means, standard deviations, and ranges of the included websites measured by the LIDA tool (the maximum achievable scores in shown in brackets after each section and question).

| Section (max. score) | Subscale | Mean (SD) | Range |
|---|---|---|---|
| **Accessibility (60)** | | 50.7 (5.3) | 37-59 |
| | Automated test (54) | 45.1 (5.4) | 31-53 |
| | Browser test (3) | 2.6 (0.7) | 1-3 |
| | Registration (3) | 3.0 (0.3) | 1-3 |
| **Usability (54)** | | 32.5 (7.1) | 19-48 |
| | Clarity (18) | 9.8 (3.2) | 3-17 |
| | Consistency (9) | 7.9 (1.5) | 3-9 |
| | Functionality (15) | 10.6 (2.4) | 6-15 |
| | Engagability (12) | 4.2 (2.6) | 1-11 |
| **Reliability (27)** | | 8.9 (5.4) | 0-22 |
| | Currency (9) | 2.0 (1.8) | 0-7 |
| | Conflicts of interest (9) | 5.6 (2.9) | 0-9 |
| | Content production (9) | 1.5 (2.2) | 0-7 |
| Total (141) | | 92.3 (13.1) | 61-127 |

**Figure 3.** Proportion of websites (N=67) categorized into three categories from poor to good, for the subscales and total LIDA score.



## Discussion

### Principal Results

We searched the Internet with different key terms to find publicly available patient information following a prenatal diagnosis of a congenital heart defect. Hits ranged from 350,000 to 67,500,000 and 67 of 600 screened websites were included in the study. Over a third (37%, 25/67) of the websites were affiliated with independent information sources, whereas a quarter (25%, 17/67) were affiliated with hospitals/clinics. The majority of the information on the websites (76%, 51/67) could not be attributed to an author. A minority of the websites contained information regarding certain prenatal aspects, that is, common feelings following prenatal diagnosis of CHD, precision of prenatal diagnosis, delivery, risks of CHD in future pregnancy, fetal intervention, termination of pregnancy, and care during pregnancy. Furthermore, the majority of the included websites were scored very poor by the DISCERN instrument with regard to information about effects of no treatment, support for shared decision making, achievement of its aims, explicit aims, risks of each treatment, how treatment choices affect overall quality of life, and areas of uncertainty. The reliability of the included websites was poor according to both the DISCERN instrument and the LIDA tool, particularly regarding currency, content production, aims of website, and areas of uncertainty. However, the accessibility and usability of the included websites were sufficient.

The literature suggests that the overwhelming number of websites found when searching the Web for information leads to information overload and searching difficulties [5,12,24]. The fact that it is difficult to find relevant information on the Internet is exemplified in this study: 39.3% (236/600) of the screened websites were irrelevant. Thus, it is possible that persons seeking information about CHD miss accurate and valuable information or give up information retrieval via the Internet because of difficulties in finding relevant sources. This difficulty might be enhanced by the fact that cardiologists seldom give recommendations on websites in connection with the diagnosis [17]. Health care professionals need to be aware that parents of children with CHD rank information regarding websites at the time of diagnosis as more important than cardiologists [25]. Consequently, as health care consumers are increasingly using the Internet to search for information [1-4], health care professionals need to address these circumstances and actively strive to recommend and provide accurate and reliable high-quality information online.

It seems that the websites target families following a postnatal diagnosis or women opting to continue the pregnancy. Previous research suggests that induced abortion is viewed as a socially unacceptable and stigmatizing procedure [26], independent of state laws on pregnancy termination [27]. It could be speculated that this perspective influenced the content of the included websites, as few contained information regarding termination as an option following a prenatal diagnosis of CHD.

The Internet may provide inaccurate and biased material [5,28,29]. It is imperative that pregnant women make informed decisions regarding whether to continue or terminate the pregnancy, which may be hindered by unreliable information sources found online. The majority of the websites in this study had poor reliability in a number of areas, including currency, conflicts of interest, and content production. The importance of current and unbiased information is especially important in the context of the rapidly expanding and evolving field of fetal cardiology, in order to promote informed decisions.

Health literacy, that is, the degree to which individuals have the capacity to obtain, process, and understand the health information and services needed to make appropriate decisions [30], is an important concept when discussing disparities in health information comprehension. Illustrations as a complement to oral information can substantially increase comprehension of health information, are especially helpful for those with poor health literacy [31], and are desired in connection with initial diagnosis [12]. However, animations and illustrations were scarce among the reviewed Web pages. It is therefore possible that the information online is not suited for those with poor health literacy. Health care professionals need to acknowledge this and provide pedagogic tools to promote patient comprehension and equal care.

### Strengths and Limitations

This study did not evaluate the scientific quality of the reviewed websites, nor did it assess the accuracy of the information found, that is, if the included websites contained any inaccurate or misleading information. Furthermore, it is possible that the key terms do not fully represent the online landscape of websites about CHD and that other results would have emerged with different search methods. However, according to previous research, the majority of health information seekers use search engines as their primary source [1,32], and the search engines used in this study are at present the most commonly used [18]. Moreover, the searches yielded 136/600 (22.7%) duplicate websites, indicating saturation and that the searches do represent the online landscape.

The DISCERN instrument and LIDA tool are based on subjective ratings. Only the first author conducted the assessments, and this could imply poor generalizability, and perhaps also a certain bias. The DISCERN instrument, developed and designed to help users of consumer health information judge the quality of written information, has been found to be consistently understood as well as transferable to different specialties (eg, [20,23,33,34]). Furthermore, the first author is a nurse, which could possibly indicate different views than non-professionals and thus different scorings. However, previous research suggests that scorings are not dependent on previous knowledge of the specific condition [22], and it has been concluded that health professionals score DISCERN similarly to non-professionals when assessing health information [33]. Taken together, we find it reasonable to assume that the main outcomes from this study would have been similar even with another evaluator or with more than one evaluator.

Approximately 25% (17/67) of the websites were affiliated with hospital/clinics, and 76% (51/67) could not be attributed to an author. Consequently, it is important to bear in mind that the information found on the included websites may differ from the information provided by health care professionals following a diagnosis of CHD.

### Suggestions for Future Research

It remains unclear if websites about congenital heart defects following a prenatal diagnosis contain accurate and suitable information. This needs to be evaluated in future studies by health professionals within fetal/pediatric cardiology and persons with experience of a prenatal diagnosis of a CHD.

It has been reported that expectant parents want more information than that provided by health care professionals following a prenatal diagnosis of CHD [17]. The results from this study, however, indicate that existing websites do not adequately supplement counseling. Development of an information source via the Internet would enable expectant parents to access accurate and tailored information that complements the standard counseling offered today. In order to evaluate such tools, national, or even international, research collaborations are needed.

Easily accessible information on the Internet influences the conditions for doctor-patient interaction [35]. What can be communicated by the doctor, and acknowledged by the patient, always depends on the previous knowledge and perspective that the patient has developed in their own information seeking. Thus, more knowledge is needed on the communication chains in which the patients build their knowledge and understanding.

The linguistic readability of the websites was not assessed in this study. Neither was patients' interpretation and evaluation of the information investigated. Several models for mechanical syntactic analysis, in order to measure readability, were developed in the early years of text linguistics, focusing mainly on factors such as word length and syntactic complexity [36,37]. Similar models have been used to assess medical information [23,34]. Later research in computational linguistics suggests that measures of semantics and discourse cohesion, that is, “content”, show higher correlations with reported readability [38]. To gain a deeper understanding of Web texts, high-quality qualitative text analytical studies are needed. Preferably, such studies should be combined with reader interviews, focusing on interpretation and comprehension. Furthermore, additional research is needed in order to understand how different types of illustrations can provide relevant understanding of the disease.

## Conclusions

The reviewed websites do not adequately supplement counseling from health care professionals following a prenatal diagnosis of CHD. Difficulties in finding relevant information sources using Web search engines and quality deficits on websites are an incentive for health professionals to take an active part in providing adequate and reliable information online about CHD. Future websites need to have a clearer prenatal perspective to become a source of knowledge for prospective parents seeking information online following diagnosis.

## Acknowledgments

This study was financed by grants from Uppsala County Council, Uppsala-Örebro Regional Research Council. Dr Bergman was also supported by Riksbankens Jubileumsfond (RJ).

## Authors' Contributions

The authors contributed to the paper as follows: TC, GB, AMK, and EM conceived and designed the study; TC collected data; TC and EM analyzed the data; TC, AMK, and EM wrote the paper; GB did a critical revision of the manuscript; TC, GB, AMK, and EM approved the final draft.

## Conflicts of Interest

None declared.

## References

1. Fox S, Duggan M. Pew Research Center. 2013 Jan 15. Health online 2013 URL: http://www.pewinternet.org/files/old-media//Files/Reports/PIP_HealthOnline.pdf [accessed 2014-05-12] [WebCite Cache ID 6PVxzeIHx]
2. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: a population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
3. Larsson M. A descriptive study of the use of the Internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]
4. Papen U. Conceptualising information literacy as social practice: a study of pregnant women's information practices. Information Research 2013 Jun;18(2):280.
5. 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]
6. Eysenbach G, Powell J, Kuss O, Sa ER. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. JAMA 2002 May;287(20):2691-2700. [Medline: 12020305]

7. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: the impact of the Internet and its implications for health care providers: findings from the first Health Information National Trends Survey. Arch Intern Med 2005 Dec;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
8. Bergman G, Borgström E, Lundell B, Sonesson SE. [Improved prenatal diagnosis of congenital heart defects. A follow-up study of prenatal ultrasound screening]. Lakartidningen 2008;105(12-13):899-903. [Medline: 18461855]
9. Marek J, Tomek V, Skovránek J, Povysilová V, Samánek M. Prenatal ultrasound screening of congenital heart disease in an unselected national population: a 21-year experience. Heart 2011 Jan;97(2):124-130. [doi: 10.1136/hrt.2010.206623] [Medline: 21163892]
10. Allan LD, Huggon IC. Counselling following a diagnosis of congenital heart disease. Prenat Diagn 2004 Dec 30;24(13):1136-1142. [doi: 10.1002/pd.1071] [Medline: 15614846]
11. Menahem S, Grimwade J. Counselling strategies in the prenatal diagnosis of major heart abnormality. Heart Lung Circ 2004 Sep;13(3):261-265. [doi: 10.1016/j.hlc.2004.06.009] [Medline: 16352205]
12. Karlsson AM, Melander Marttala U, Mattsson E. [We can not show a normal heart: Communicative praxis in connection to heart defects in fetus]. Svenskans beskrivning 2014;33:193-213.
13. Sharland G. Fetal cardiac screening: why bother? Arch Dis Child Fetal Neonatal Ed 2010 Jan;95(1):F64-F68. [doi: 10.1136/adc.2008.151225] [Medline: 20019198]
14. Vandvik IH, Førde R. Ethical issues in parental decision-making. An interview study of mothers of children with hypoplastic left heart syndrome. Acta Paediatr 2000 Sep;89(9):1129-1133. [Medline: 11071097]
15. Wool C. Systematic review of the literature: parental outcomes after diagnosis of fetal anomaly. Adv Neonatal Care 2011 Jun;11(3):182-192. [doi: 10.1097/ANC.0b013e31821bd92d] [Medline: 21730912]
16. Brosig CL, Whitstone BN, Frommelt MA, Frisbee SJ, Leuthner SR. Psychological distress in parents of children with severe congenital heart disease: the impact of prenatal versus postnatal diagnosis. J Perinatol 2007 Nov;27(11):687-692. [doi: 10.1038/sj.jp.7211807] [Medline: 17717519]
17. Hilton-Kamm D, Sklansky M, Chang RK. How not to tell parents about their child's new diagnosis of congenital heart disease: an Internet survey of 841 parents. Pediatr Cardiol 2014 Feb;35(2):239-252. [doi: 10.1007/s00246-013-0765-6] [Medline: 23925415]
18. eBizMBA. 2014. Top 15 most popular search engines: May 2014 URL: http://www.ebizmba.com/articles/search-engines [accessed 2014-05-12] [WebCite Cache ID 6PVzUmkSd]
19. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
20. Charnock D. Discern online.: University of Oxford and The British Library; 1998. The DISCERN handbook: quality criteria for consumer health information on treatment choices URL: http://www.discern.org.uk/discern.pdf [accessed 2014-05-12] [WebCite Cache ID 6PVzbZzIL]
21. Minervation. 2007. The LIDA instrument: minervation validation instrument for health care web sites URL: http://www.minervation.com/wp-content/uploads/2011/04/Minervation-LIDA-instrument-v1-2.pdf [accessed 2014-05-12] [WebCite Cache ID 6PVzhVPIO]
22. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999 Feb;53(2):105-111 [FREE Full text] [Medline: 10396471]
23. Grewal P, Alagaratnam S. The quality and readability of colorectal cancer information on the internet. Int J Surg 2013 Mar;11(5):410-413. [doi: 10.1016/j.ijsu.2013.03.006] [Medline: 23523948]
24. Wathen CN, Harris RM. "I try to take care of it myself." how rural women search for health information. Qual Health Res 2007 May;17(5):639-651. [doi: 10.1177/1049732307301236] [Medline: 17478646]
25. Arya B, Glickstein JS, Levasseur SM, Williams IA. Parents of children with congenital heart disease prefer more information than cardiologists provide. Congenit Heart Dis 2013 Feb;8(1):78-85 [FREE Full text] [doi: 10.1111/j.1747-0803.2012.00706.x] [Medline: 22891764]
26. Astbury-Ward E, Parry O, Carnwell R. Stigma, abortion, and disclosure--findings from a qualitative study. J Sex Med 2012 Dec;9(12):3137-3147. [doi: 10.1111/j.1743-6109.2011.02604.x] [Medline: 22239919]
27. Shellenberg KM, Moore AM, Bankole A, Juarez F, Omideyi AK, Palomino N, et al. Social stigma and disclosure about induced abortion: results from an exploratory study. Glob Public Health 2011 Jul;6 Suppl 1:S111-S125. [doi: 10.1080/17441692.2011.594072] [Medline: 21745033]
28. Cooke A. Quality of health and medical information on the Internet. Clin Perform Qual Health Care 1999 Dec;7(4):178-185. [Medline: 10947392]
29. Resnik DB. Patient access to medical information in the computer age: ethical concerns and issues. Camb Q Healthc Ethics 2001;10(2):147-54; discussion 154-6. [Medline: 11302091]
30. Nielsen-Bohlman L, Panzer AM, Hamlin B, Kindig DA. Health literacy: a prescription to end confusion. Washington, DC: National Academies Press; 2004.

31. Houts PS, Doak CC, Doak LG, Loscalzo MJ. The role of pictures in improving health communication: a review of research on attention, comprehension, recall, and adherence. Patient Educ Couns 2006 May;61(2):173-190. [doi: 10.1016/j.pec.2005.05.004] [Medline: 16122896]
32. Harris interactive. 2011. The Growing Influence and Use Of Health Care Information Obtained Online URL: http://www.harrisinteractive.com/vault/HI-Harris-Poll-Cyberchondriacs-2011-09-15.pdf [accessed 2014-05-12] [WebCite Cache ID 6PW0VrtSp]
33. Griffiths KM, Christensen H. Website quality indicators for consumers. J Med Internet Res 2005;7(5):e55 [FREE Full text] [doi: 10.2196/jmir.7.5.e55] [Medline: 16403719]
34. McKearney TC, McKearney RM. The quality and accuracy of internet information on the subject of ear tubes. Int J Pediatr Otorhinolaryngol 2013 Jun;77(6):894-897. [doi: 10.1016/j.ijporl.2013.03.021] [Medline: 23587675]
35. Freebody P, Freiberg J. Health literacy as a social practice: response to Nutbeam. Literacy and Numeracy Studies 1999;9(2):57-66.
36. Björnsson CH. [Readability]. Stockholm: Liber; 1968.
37. Kincaid J, Fishburne R, Rogers R, Chissom B. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel. Tech Repp 1975;8:e75.
38. Pitler E, Nenkova A. Revisiting readability: a unified framework for predicting text quality. 2008 Presented at: Proceedings of the Conference on empirical methods in natural language processing; 2008; Honolulu, Hawaii p. 186-195.

## Abbreviations

**CHD:** congenital heart defect

*Edited by G Eysenbach; submitted 30.08.14; peer-reviewed by C Bermúdez-Tamayo, K Friedman; comments to author 29.10.14; accepted 23.11.14; published 21.01.15.*

*Please cite as:*
*Carlsson T, Bergman G, Karlsson AM, Mattsson E*
*Content and Quality of Information Websites About Congenital Heart Defects Following a Prenatal Diagnosis*
*Interact J Med Res 2015;4(1):e4*
*URL: http://www.i-jmr.org/2015/1/e4/*
*doi:10.2196/ijmr.3819*
*PMID:25608457*

Short Paper

# Patients' Perspectives of Accessibility and Digital Delivery of Factual Content Provided by Official Medical and Surgical Specialty Society Websites: A Qualitative Assessment

Darren Ow[1], MBBS; David Wetherell[1*], MBChB; Nathan Papa[1], MBBS; Damien Bolton[1*], MBBS, MD, FRACS; Nathan Lawrentschuk[1,2,3], MBBS, PhD, FRACS

[1]Austin Hospital, Department of Surgery, University of Melbourne, Melbourne, Australia
[2]Olivia Newton-John Cancer Research Institute, Austin Hospital, Melbourne, Australia
[3]Peter MacCallum Cancer Centre, Department of Surgical Oncology, Melbourne, Australia
[*]these authors contributed equally

**Corresponding Author:**
Darren Ow, MBBS
Austin Hospital
Department of Surgery
University of Melbourne
Urology Office, Level 8 Harold Stokes Building, Austin Hospital
145 Studley Road, Heidelberg
Melbourne, 3084
Australia
Phone: 61 412813240
Fax: 61 394963617
Email: okhdarren@gmail.com

## Abstract

**Background:** Health care websites provide a valuable resource of health information to online consumers, especially patients. Official surgical and medical society websites should be a reliable first point of contact.

**Objective:** The primary aim of this study was to quantitatively assess medical and surgical society websites for content and highlight the essential features required for a high-quality, user-friendly society website.

**Methods:** Twenty specialty association websites from each of the regions, Australia, UK, Canada, Europe, and the USA were selected for a total of 100 websites. Medical and surgical specialities were consistent across each region. Each website was systematically and critically analysed for content and usability.

**Results:** The average points scored per website was 3.2 out of 10. Of the total (N=100) websites, 12 scored at least 7 out of 10 points and 2 scored 9 out of 10. As well, 35% (35.0/100) of the websites had an information tab for patients on their respective homepages while 38% (38.0/100) had download access to patient information. A minority of the websites included different forms of multimedia such as pictures and diagrams (24.0/100, 24%) and videos (18.0/100, 18%).

**Conclusions:** We found that most society websites did not meet an adequate standard for delivery of information. Half of the websites were not patient accessible, with the primary focus being for health professionals. As well, most required logins for information access. Specialty health care societies should create patient-friendly websites that would be beneficial to all online consumers.

***(Interact J Med Res 2015;4(1):e7)*** doi:10.2196/ijmr.3963

**KEYWORDS**

Internet; online health information; patient education; social media; Web science

## Introduction

Advances in modern technology and communication have resulted in a new digital age where the Internet is an important source of information for health professionals and consumers. The number of health care providers and consumers gaining access to this information is expanding [1]. The Pew Internet and American Life Project reported that 80% of adults in the United States of America (USA) sought health information via search engines (Google, Bing, or Yahoo) [2]. Many specialities have embraced this technological advance. For example, in recent years, urology has been open to integrating the Internet and social media as a new communication platform [3]. Although social media has been around for the past decade and is widely used in other spheres, it has not been utilised as much in the health community, and only a small number of health faculties engage in social media [4]. Additionally, there is a wide disparity in the quality of health information on the Internet and not all accredited and quality information is readily accessible for the online consumer.

Official surgical and medical society websites are a valuable resource of health information for professionals and patients. They allow for the centralisation of information in a user-friendly and accessible format. Within this niche of websites, the quality of information remains variable. Our primary aim was to systematically assess these websites for content, quality, and delivery of health information. We also intended to highlight the key features required for a high-quality society website.

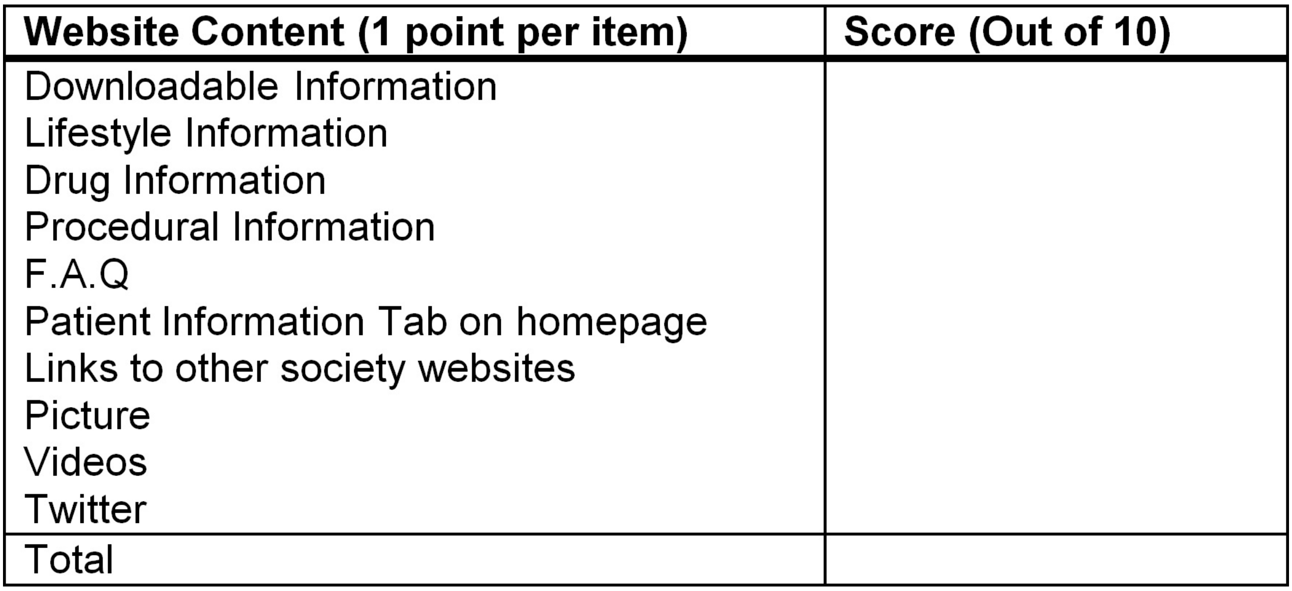
## Methods

Medical and surgical society websites (N=100) were systematically reviewed in September 2012. Websites (n=20) from each of the regions (Australia, Canada, Europe, United Kingdom, and USA) were collected from various medical (cardiology, endocrinology, gastroenterology, haematology, infectious disease, nephrology, neurology, oncology, respiratory and rheumatology) and surgical (cardiothoracic, general surgery, maxillofacial, neurosurgery, otolaryngology, paediatric surgery, plastics, urology and vascular) specialties. A Google search was conducted to identify the websites using the keywords medical or surgical specialty name, society or association, and country. All selected websites were in English.

Rafe et al constructed a qualitative framework to assess hospital and other medical websites that focused on 7 key metrics, which are (1) content, (2) design, (3) organization, (4) user-friendliness, (5) performance, (6) service, and (7) technical quality [5]. Using this framework of content and user-friendliness metrics, we constructed a simple 10-point quality appraisal tool. This tool was designed to assess the usability of websites for health professionals, patients, and other online consumers. The 10 points included information about procedures, drugs, and lifestyle interventions, a frequently asked questions (FAQs) page, pictures and diagrams, video attachments, social media links (ie, Twitter), presence of a patient information tab on the homepage, ability to easily download information, and inclusion of relevant website links (Figure 1).

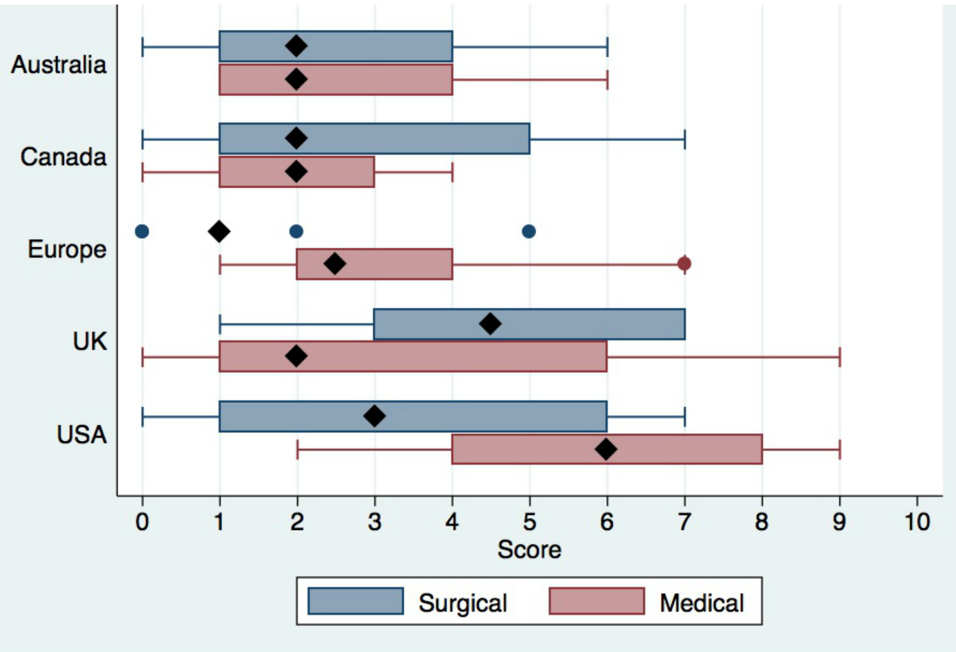
The variables were evaluated for statistically significant differences between regions and specialties using the non-parametric Kruskal-Wallis equality of population test. Statistical significance was set at *P*<.05. The scores were graphed using a box and whisker plot (Figure 2). Statistical analysis was performed using Stata v.12.0 SE (College Station, Texas).

**Figure 1.** Website content and usability scoring sheet.



| Website Content (1 point per item) | Score (Out of 10) |
|---|---|
| Downloadable Information<br>Lifestyle Information<br>Drug Information<br>Procedural Information<br>F.A.Q<br>Patient Information Tab on homepage<br>Links to other society websites<br>Picture<br>Videos<br>Twitter | |
| Total | |

**Figure 2.** Webscores based on country and specialty.



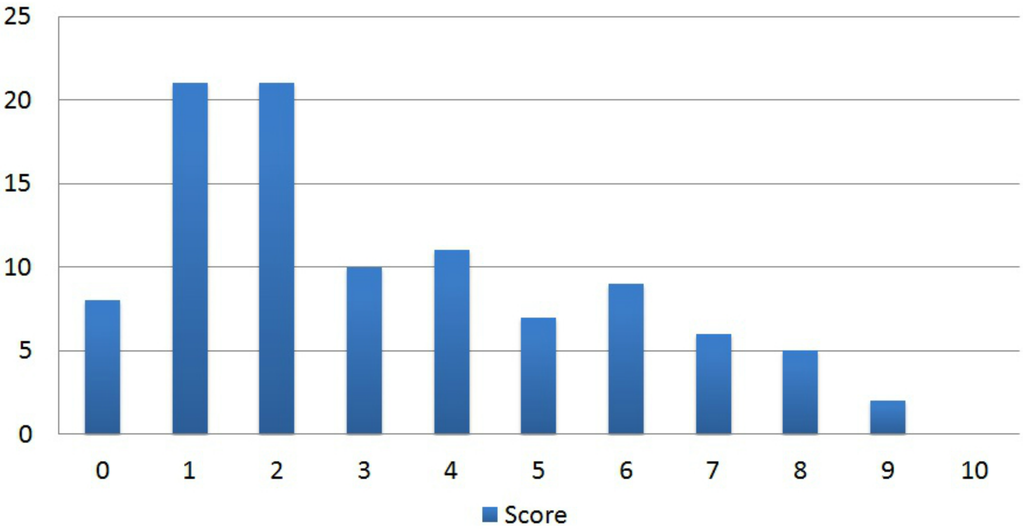
## Results

A total of 100 health care society websites were selected in this study with an equal number of medical and surgical specialties. The mean points scored per website was 3.2 (range 0-9, SD 2.4) and the median points per site was 2.5. Of the websites assessed, 12 scored at least 7 points, 2 scored 9 points, and 28 scored between 4-6 points out of a total of 10. The remaining 60% (60.0/100) of websites scored 3 or less points with 9 scoring 0 points (Figure 3).

Among the 5 regions, USA health care websites had the highest mean website score of 4.6, followed by UK websites with a mean score of 4.0. Australian and Canadian health care society websites scored 2.7 and 2.6, respectively, while the European health care society websites scored a mean of 2.2 points. Statistical testing revealed a significant difference between the regions (*P*=.01). When comparing medical and surgical specialties, surgical specialties had a mean score of 2.9 points per website, while medical specialties had a mean score of 3.4 points. These scores were not statistically significant.

Of the 100 society websites, 35 (14 medical, 21 surgical) displayed a patient information tab on their main webpage. Links to other health care related websites were displayed on 60 websites. A minority of websites included different forms of multimedia such as pictures and diagrams (24.0/100, 24%) and videos (18.0/100, 18%). As well, 38 websites (15 medical, 23 surgical) had information that was easily downloadable without any login requirements, 27 (15 medical, 12 surgical) had lifestyle intervention information, 21 (14 medical, 7 surgical) had drug information, 36 (14 medical, 22 surgical) had procedural information, and 7 had a FAQs page. Out of 100 associations, 61 (61.0/ 61%) had active Twitter accounts and of these, 52 (85.0/100, 85%) had direct links to these accounts from their websites.

**Figure 3.** Total website scores based on points.



## Discussion

### Principal Findings

Our website quality appraisal tool was initially designed to assess the accessibility of health care society websites from a patient's perspective. Almost all medical and surgical societies have established official websites. These websites should be the first point of reference for online health information relevant to a particular condition covered by that specialty. However, these websites achieved a wide range of scores when the quality appraisal tool was applied. Only a few health care society websites managed to score 7 or more points while most failed to display adequate content and usability, thus making it difficult for patients and health care professionals to use the websites as a reliable information source.

We found that about one third of these official websites had patient information tabs in their homepage with the incentive being for patients to gain a better understanding of their own medical condition or about the specialty. A number of websites were designed with pictures (n=24) and videos (n=18) attached to the relevant conditions discussed [6].

Other forms of multimedia such as podcasts and power point lectures were provided more for health care professionals. In terms of drug information being displayed on websites, medical societies had higher scores than surgical societies, while the reverse was found with respect to procedural information. This might be explained by the fact that medical specialties make more use of treatments with medication rather than surgical procedures.

Interestingly we observed that website scores varied for each region. Both the USA and UK society websites had higher scores than other regions. Clearly, these websites showed a more user-friendly approach towards patients than Australian, Canadian, and European society websites. European health care society websites scored the lowest among the other regions, explained by the paucity of the available written and multimedia information. However, this may also be due to the possibility of alternative websites in other languages.

One possible reason the USA and UK websites scored highly is that at the time of review, about 80% (80.0/100) of USA and UK health care society websites had links to official Twitter accounts, while only 20% (20.0/100) of the Australian and Canadian society websites had an official Twitter account link. Some health care societies have used Twitter as a platform to engage with the public community or even with health care professionals [7]. Twitter has been utilised at many conferences to engage in clinical discussions and to further the communication outside of the conference [8,9]. Despite the exponential increase of Twitter use in the medical field over the past few years, these forms of social media are still not widely accepted in certain countries [10].

There is a proliferation of medical information websites on the Internet, most of uncertain quality. As an official society website, the health information provided therein is considered reliable and accurate. However, websites requiring a login to access online information or other features may deter patients from exploring the website any further. A possible explanation for restricted access could be that the majority of the websites were designed for health care professionals rather than patients. It would be ideal if this information was more widely disseminated to all health consumers [11]. Free access to online information could potentially be an alternative method to deliver information and improve communications between health care professionals and patients, thus narrowing the gap to health care services [12].

### Limitations

A limitation to the study was that some of the medical societies had more than one official website and therefore, there was a lack of centralisation of information. Therefore, it is our opinion that society-association websites should be unified, although this may not be feasible in certain circumstances where websites have already been established.

The appraisal tool used in this study was designed towards a patient-focused evaluation rather than health care professional. Therefore, other potential key roles of society websites that were designed for health care have not been evaluated. Also, this study only included English language websites and may be biased towards regions where English is not the first language.

The presence of Twitter links to each health care society websites were reviewed, however, the level of Twitter activity was not assessed. Twitter activity could potentially be a better indicator of engagement with online consumers.

### Conclusions

In this era, the population of online users seeking health information has risen, and health care societies should try to adapt and make the transition to developing a higher standard of website. This will engage and encourage patients to participate in their health management rather than being a passive recipient of health care. Creating good quality websites by using a patient-friendly framework would be beneficial to all online consumers. However, most society websites were created specifically for health care professionals rather than patients. Furthermore, these health care society websites were often not user-friendly for the patient with some requiring a login for information access. Although specialty health care societies-associations' websites have been established for health care professionals, these websites need to become more education-focused for patients if they are to be the lead voice in their area of practice, and to improve their craft's profile within the wider community.

### Conflicts of Interest

None declared.

## References

1. Jadad AR, Gagliardi A. Rating Health Information on the Internet. JAMA 1998 Feb 25;279(8):611-614. [doi: 10.1001/jama.279.8.611] [Medline: 9486757]
2. Fox S, Duggan M. 2013 Nov 23. Part Two: Sources of Health Information URL: http://www.webcitation.org/6Ww8RKMxw [accessed 2015-03-10]
3. Keller B, Labrique A, Jain KM, Pekosz A, Levine O. Mind the gap: social media engagement by public health researchers. J Med Internet Res 2014;16(1):e8 [FREE Full text] [doi: 10.2196/jmir.2982] [Medline: 24425670]
4. Matta R, Doiron C, Leveridge MJ. The dramatic increase in social media in urology. J Urol 2014 Aug;192(2):494-498. [doi: 10.1016/j.juro.2014.02.043] [Medline: 24576656]
5. Rafe V, Monfaredzadeh M. A qualitative framework to assess hospital / medical websites. J Med Syst 2012 Oct;36(5):2927-2939. [doi: 10.1007/s10916-011-9771-5] [Medline: 21874337]
6. Bollschweiler E, Apitzsch J, Apitsch J, Obliers R, Koerfer A, Mönig SP, et al. Improving informed consent of surgical patients using a multimedia-based program? Results of a prospective randomized multicenter study of patients before cholecystectomy. Ann Surg 2008 Aug;248(2):205-211. [doi: 10.1097/SLA.0b013e318180a3a7] [Medline: 18650629]
7. Griffis HM, Kilaru AS, Werner RM, Asch DA, Hershey JC, Hill S, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. J Med Internet Res 2014;16(11):e264 [FREE Full text] [doi: 10.2196/jmir.3758] [Medline: 25431831]
8. Wilkinson SE, Basto MY, Perovic G, Lawrentschuk N, Murphy DG. The social media revolution is changing the conference experience: analytics and trends from eight international meetings. BJU Int 2014 Aug 18. [doi: 10.1111/bju.12910] [Medline: 25130687]
9. Nason GJ, O'Kelly F, Kelly ME, Phelan N, Manecksha RP, Lawrentschuk N, et al. The emerging use of Twitter by urological journals. BJU Int 2015 Mar;115(3):486-490. [doi: 10.1111/bju.12840] [Medline: 24925047]
10. Brown J, Ryan C, Harris A. How doctors view and use social media: a national survey. J Med Internet Res 2014;16(12):e267 [FREE Full text] [doi: 10.2196/jmir.3589] [Medline: 25470407]
11. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
12. Cadogan M, Thoma B, Chan TM, Lin M. Free Open Access Meducation (FOAM): the rise of emergency medicine and critical care blogs and podcasts (2002-2013). Emerg Med J 2014 Oct;31(e1):e76-e77. [doi: 10.1136/emermed-2013-203502] [Medline: 24554447]

---

## Abbreviations

**FAQs:** frequently asked questions

---

*Edited by G Eysenbach; submitted 23.10.14; peer-reviewed by J Ischia, M Leveridge; comments to author 04.12.14; revised version received 21.01.15; accepted 07.02.15; published 27.03.15.*

*Please cite as:*
*Ow D, Wetherell D, Papa N, Bolton D, Lawrentschuk N*
*Patients' Perspectives of Accessibility and Digital Delivery of Factual Content Provided by Official Medical and Surgical Specialty Society Websites: A Qualitative Assessment*
*Interact J Med Res 2015;4(1):e7*
*URL: http://www.i-jmr.org/2015/1/e7/*
*doi:10.2196/ijmr.3963*
*PMID:25830824*

Original Paper

# Medical Information on the Internet: A Tool for Measuring Consumer Perception of Quality Aspects

Arthur Dubowicz[1,2*], PhD; Peter J Schulz[1*], PhD

[1]Institute of Communication and Health, University of Lugano, Lugano, Switzerland

[2]Department of Society and Human Development, Harvard School of Public Health, Boston, MA, United States

[*]all authors contributed equally

**Corresponding Author:**
Arthur Dubowicz, PhD
Institute of Communication and Health
University of Lugano
Via G. Buffi 13
Lugano, 6904
Switzerland
Phone: 41 58 666 48 21
Fax: 41 58 666 46 47
Email: arthur.dubowicz@usi.ch

## Abstract

**Background:** Most of adult Internet users have searched for health information on the Internet. The Internet has become one of the most important sources for health information and treatment advice. In most cases, the information found is not verified with a medical doctor, but judged by the “online-diagnosers” independently. Facing this situation, public health authorities raise concern over the quality of medical information laypersons can find on the Internet.

**Objective:** The objective of the study was aimed at developing a measure to evaluate the credibility of websites that offer medical advice and information. The measure was tested in a quasi-experimental study on two sleeping-disorder websites of different quality.

**Methods:** There were 45 survey items for rating the credibility of websites that were tested in a quasi-experimental study with a random assignment of 454 participants to either a high- or a low-quality website exposure. Using principal component analysis, the original items were reduced to 13 and sorted into the factors: trustworthiness, textual deficits of the content, interferences (external links on the Web site), and advertisements. The first two factors focus more on the provided content itself, while the other two describe the embedding of the content into the website. The 45 survey items had been designed previously using exploratory observations and literature research.

**Results:** The final scale showed adequate power and reliability for all factors. The loadings of the principal component analysis ranged satisfactorily (.644 to .854). Significant differences at $P<.001$ were found between the low- and high-quality groups. Advertisements on the website were rated as disturbing in both experimental conditions, meaning that they do not differentiate between good and bad information.

**Conclusions:** The scale reliably distinguished high- and low-quality of medical advice given on websites.

***(Interact J Med Res 2015;4(1):e8)*** doi:10.2196/ijmr.3144

**KEYWORDS**

online health information; scale development; quality assessment; sleeping disorders; Internet intervention

## Introduction

### Health Information and the Internet

Internet usage is increasing strongly as more and more people have access to it. The increase reaches all age groups, including older people [1,2]. As a result, the amount and the use of health-related information on the Internet are also growing. Several studies show that, for health information, the Internet is one of the primary resources [3-7]. The Internet has thus become one of the most important sources for health information and for searching health care services and treatment advice. Data show that, within a given year, about 80% of adult Internet users have searched for health information [3,8]. In Germany in 2007, 56.6% of Internet users described their use as health-related [9]. In comparison to a previous study, Germany was among the European countries with the highest growth in this segment [9]. The age group searching most actively for health information was young adults between the ages of 30 and 44 years [10]. Data also show that, with a higher usage and availability of the Internet in general, Internet health usage grew across all age groups and among both genders [9]. About 35% of people searching for health information use the information they find to diagnose their medical condition. Only half of these so-called “online-diagnosers” check their diagnosis with a medical professional [3].

Cost and time factors make searching the Internet an attractive alternative to seeing a doctor in a nonacute situation, as information is available immediately and a visit to one’s doctor can be (work) “time consuming”. Individual reasons for searching medical information might differ—some want to prepare for a medical doctoral consultation, others seek support, or alternative remedies to treatment advice—but the accuracy of search results is significant for “online-diagnosers”. Hence, public health authorities are concerned over the quality of the health information available on the Internet [10]. A review on mental disorder information websites came to the conclusion that most scholarly articles report poor quality [11]. Erroneous, misleading, or irrelevant health information provided on the Internet can lead to wrong self-diagnosis and ineffectual or damaging treatment attempts by the layperson, and to delayed presentation at a general practitioner or hospital, which in turn can make therapy more difficult. This risk is especially increased by the fact that most of the information found on the Internet is not discussed with a medical practitioner, but rather used as the single basis for making a decision [8,12]. In addition, information acquired from the Web might make patients less willing to adhere to their doctor’s advice, and thus result in poor health outcomes. Finally, there is also the possibility of financial damages if a patient decides, based on bad advice from websites, to buy over-the-counter medication or equipment that does not provide remedy. Health-related decisions of individuals can be understood as affected by health literacy, which is the ability to understand medical information and to pass adequate judgment in matters of health [13,14]. The ability to distinguish good advice from bad advice can therefore be considered an aspect of health literacy.

### Sleeping Disorders

A very common medical condition in the general population is sleeping disorders or insomnia. About 50% of the population complains about such problems in a given year, and it is the most common complaint of patients after general pain [15]. Moreover, most people suffer from sleeping disorders periodically, and often have to rely on self-treatment when not at a doctors [16-18]. A lot of information on this condition can be found on the Internet. This material is very diverse, and the corpus consists of medical information, individual reports, advertisements, as well as alternative remedies. Moreover, producers of over-the-counter sleeping medication advertise their products heavily. As there is so much and such diverse information, its quality becomes difficult to judge. Additional research has shown that the accuracy of health information depends on the topic; information on more specific diseases is of higher quality than information on general health problems [19]. Especially in this context, sleeping disorders can be seen as a condition with much low-quality information. For this study, the whole range of sleeping disorders was incorporated, and no selection was applied as to whether it was a primary disease or a symptom.

### Credibility of Internet Health Information

The understanding of trust and credibility factors of Internet health information, and websites in general, has been addressed by research in recent years. Accordingly, various measures and quality criteria for health information on the Internet can be found [19-28]. An often-found approach is based on expert or consumer ratings of health information [25,29-31]. The DISCERN scale and its adaptation for the eHealth context are the prime examples, assessing health information quality with regard to patients’ treatment decisions [32,33]. In contrast, our measure tries to take the particular setting of Internet information into consideration. The DISCERN scale was developed for health or communication professionals and experienced users who want to discriminate between high- and low-quality health information. In contrast, our approach tried to take the particular setting of Internet health information into consideration and puts the average user of health information into its focus [33]. Information usage on the Internet is characterized by the short attention given by the consumers and a comparison of different sources [3,5,7].

A recent review described some of the tools for assessing the quality as having limited validity [11]. Still, most of these tools lack empirical testing and provide mostly conceptual work [28]. Reviews in the field mention the lack of an overall framework to assess this domain, and the need for a feasible definition of quality criteria for the websites [19]. There is also research on the process of how consumers assess medical information on websites [34]. Another line of research is focusing more strongly on the factors which make a website with health information a credible source for consumers [35]. Whether Internet health information consumers are able to determine the quality of the information found remains unanswered.

Another line of research assesses quality aspects of health information websites through predefined key word lists evaluating the provided metadata of websites [36]. These measures often combine a checklist for health-relevant words with cross checks of different websites in this domain [37]. Still, these approaches focus on information provided by the hosting provider or institution responsible for the Internet information. Additionally, the provided content is often analyzed for readability and difficult wording [38]. In contrast to these approaches, the aim of this research is to investigate the ability of individuals to distinguish the quality of health information websites. A measure was designed within the context of German language health information on the Internet. The medical condition of sleeping disorders or insomnia was chosen. The procedure for developing and evaluating this measure followed mostly the structured theoretical approach of DeVells [39]. Adaptations were made when combining qualitative and quantitative methods for including the consumer's perspective, and due to Internet-specific data collection techniques. For the development, observations and structured post observation interviews were used. Based on the findings, a measure was designed. It was tested with a two-group experimental analysis in an Internet survey.

## Methods

### Preliminary Observational Study

To evaluate how a Web search is conducted, 42 naturalistic observations of individuals searching the Internet for information on sleeping disorders were collected. The participants were asked to search for information about sleeping disorders in general; the search was not limited to a distinct perspective or a certain type of sleeping disorder. Following the individual search on the Internet, post observational, structured, in-depth interviews were conducted to clarify users' motivation for particular search decisions and obtain additional information on their search behavior.

Undergraduate students were instructed to contact volunteer participants in their neighborhood and to observe their searching behavior. The observers were instructed following the guidelines of DeWalt and DeWalt [40]. Particular focus was given to actively observing and taking note of details which would be relevant for the protocols, taking note of possible uncertainties or difficulties of the participants to be clarified in the follow-up interview [40]. Most students contacted the participants in coffee shops where Wi-Fi was available and laptops were being used. To approach them, the student observers were equipped with an observation sheet and interview protocols. The participants received a short study objective beforehand. Participants were informed that this observation was conducted by university students for a research project on health information on the Internet.

The research group designed a field protocol for this study in order to capture the observed setting and contents, following previous recommendations of Schensul et al [41]. The protocols allowed registering the participants' sociodemographics, the search procedure, the exact search term, their selection from a search results list, the length of time they remained on a website, and the number of results they opened within the observation period. These observation protocols were discussed later in an interview with the participants to collect additional information on their reasons for their choices during the search. In addition, the participants were asked for aspects they remembered from the visited websites. According to Bogdan and Biklin, process codes and activity codes were used to study the participants' search strategies as described in the protocols [42]. The observation protocols were analyzed following the search procedure of the participants. Similarities and outliers were found by identifying the codes on conferring content equivalence and according to the statements given by the participants.

### Measure Development

Based on the conclusions of the observational study and the interviews, a multi-item measure for the credibility of health websites on the Internet was designed. Orientation for this study was found in the previous work on measures of health information quality assessment [3,7,8,34] and literature reviews in this field [19]. The procedure led to a scale consisting of seven dimensions, each composed of several items, summing up to 49 items in total. The items are designed in the format of statements to which participants can concur or oppose on a seven-point scale ranging from 1 "completely disagree" to 7 "completely agree". This preliminary scale was critically discussed within the research group, taking the literature into account. Moreover, the single items were checked and pretested with 14 undergraduate students. If necessary, they were adapted, leading to the final measure consisting of seven dimensions and 45 items in total. The dimensions cover several aspects, which were, in the preliminary study, identified as relevant. Among them are more general dimensions (such as layout of the website, textual deficits, usability, and interferences due to advertisement banners and others) and more content-oriented dimensions (such as a trustworthy source, the competence of the authors, and the suitability of the given information for everyday life). The dimensions and numbers of items are presented in Table 1. The items that compose the final measure are shown in Table 4.

**Table 1.** The dimensions of the measure based on the observational study.

| Dimension | Number of items | Interest |
|---|---|---|
| Trustworthiness | 8 | Trustworthy source |
| Competence | 7 | Content is adequate |
| Interference | 7 | Pop-up windows, advertisement |
| Layout | 7 | Presentation style |
| Textual deficits | 8 | Factor of intelligibility |
| Usability | 4 | Access to the information |
| Suitability | 4 | Implementation of the advice |

### Implementation of the Internet Survey

To test the developed scale, an Internet survey was designed, comparing a group exposed to a low-quality site with another one exposed to a high-quality site. Participants were recruited in two weeks through a snowball system via email, social networks, and online-communities. It was initiated with a sample of 14 undergraduate students. The participants were randomly assigned to one of the two conditions. The high-quality website was rated as such by an independent German consumer foundation involved in investigating and comparing goods and services in an unbiased way [43]. The other website was rated as having low-quality content by the research group in collaboration with sleep experts. For both websites, standardized readability formulas were used to calculate the general reading level. Both websites were of medium complexity. The high-quality website scored 52.61, while the low-quality website scored 47.35 on a scale from 0 (easiest) to 100 (most difficult) [44]. The content of both websites was checked for quality. The key elements were accuracy of the medical information provided, ease of navigation on the website, moderation by the provider, structure and style of content, and if an advertisement could be easily recognized as such. The content of the low-quality website was based on a very general description of insomnia symptoms. Moreover, no sources for the given information were mentioned, which is why it was not clear whether experts or expert knowledge were involved in producing the written content. User comments were neither sorted nor reviewed. Furthermore, it was difficult to distinguish between links for auxiliary insomnia-related content and insomnia-related advertisements. Both were general health websites; only the sections about sleeping disorders were the subject of investigation. For embedding the websites, a HyperText Markup Language (HTML) snippet with the technical restrictions was included into the Internet survey software. HTML is a commonly used markup language for designing Web pages. This Internet survey was administered by a noncommercial and university-based Internet survey platform. Such procedure was inspired by the possibilities offered through digital media and the widely used combinations of research design in offline surveys.

The Internet survey incorporated the websites, and participants had to explore the content for at least four minutes; otherwise it was not possible to continue. The interfaces of the websites were included into the Internet survey mask, while external links on the websites were blocked. Internal paths leading away from sleeping-disorder content were blocked. The quality certificates shown on the high-quality website were removed. The survey was technically pretested before being distributed. After the website exposure, the Internet survey started. The 45 items of the credibility scale and the four items of the outcome measure were presented to each participant in a different random order. At the end of the survey, the participants were asked to respond to questions regarding their Internet usage of health information sites, occupation in a medical profession, and sociodemographic information.

To measure the impact of the website on participants' behavior, an outcome measure was added. It consisted of four items formulating future intention to consult the site, intention to recommend it, etc (Textbox 1). To achieve a single measure, the items were later averaged. They had the same scaling as the 45 credibility items and were asked in random order together with them.

**Textbox 1.** The items of the outcome measure as used in the Internet survey.

Outcome measure:

I would recommend this website

I would approach this source for future questions

I can trust the information on this website

If I suffered from sleeping-disorders, I would use the given information

### Data Analysis

To assess the internal consistency of the measure, a scale reliability analysis was conducted. To check for differences between sociodemographic groups and occupations, respectively, Internet usage for searching medical information, correlations was used. For reasons of sound data analysis, the negatively worded items were reversed using the formula NEWSCORE= (MAX + MIN) – SCORE.

Factors were identified when in the simple structure approach eigenvalues greater than 1.0 were computed [45]. An adequate sample size was checked, using a ratio of five cases to one variable. Following the methodological recommendation presented by Gerbing and Hamilton [46], first a principal component analysis using the Kaiser Normalization and a Varimax rotation was conducted. Moreover a Promax rotation for the identified factors was computed to check their correlations. For the measure of sampling adequacy, factor loadings below .5 were excluded [47]. For all computations, an IBM SPSS Statistics 21 software package was used [48].

## Results

### Observational Study

The participants of the observational study (N=42) were mainly male (25/42, 60%), between 21 and 40 years old, and most had some university degree (20/42, 48%). Table 2 provides a detailed description of these characteristics. The search time was limited to ten minutes by the observers. For the follow-up interviews, between five and ten minutes were needed.

When searching for information on sleeping disorders, all participants used the "Google" search engine as a starting point. Other portals or direct access to websites of medical authorities were not considered. This seems to be in accordance with other recent findings [3,34,49,50]. While some participants were searching for the terms "sleeping disorders" others added a "condition related" term such as "treatment" or "help". Some participants were very effective in combining these search terms or also using Boolean combinations and sign operators; those that did not had more difficulty finding adequate results, which took more time spent in checking the result list and deciding which website to choose. There were ten participants that exclusively opened results that were displayed on the first result page of the search engine. A page showed a list with ten results; to see more results, participants had to navigate to the next result page. None of the participants checked more than six result pages. Previous research on search behavior notes that the first results are the most likely to be looked at [34]. Sponsored links shown before the results were not taken into consideration in the participants' search.

In the interviews, the participants were asked individually about their personal observation protocol. They reported that the most relevant key factor for choosing a specific website was its name. The observation protocols showed that a simpler domain name is more likely to be clicked, especially if the search-term was an integral part of the name. As reasons for staying on a website and checking the provided information, most participants mentioned a friendly layout and quality content. Commonly mentioned reasons for leaving were disturbances by advertisement or pop-up boxes and nonadequate information (too general or too specific). About 15 participants stressed the importance of a credible author, such as a governmental institution, a medical association, or professional medical personal, as factors to open or stay on a website.

**Table 2.** Detailed sample description of the observational study, N=42.

| Participants | | n | % |
|---|---|---|---|
| Total number, N | | 42 | 100 |
| **Gender** | | | |
| | Male | 25 | 60 |
| | Female | 17 | 41 |
| **Age group, years** | | | |
| | 17-20 | 5 | 12 |
| | 21-30 | 10 | 24 |
| | 31-40 | 12 | 29 |
| | 41-50 | 8 | 19 |
| | 50-62 | 7 | 17 |
| **Education** | | | |
| | No school degree | 1 | 2 |
| | Some school degree | 7 | 17 |
| | High school degree | 5 | 12 |
| | Professional school degree | 5 | 12 |
| | In university education | 4 | 10 |
| | University degree | 20 | 48 |

## Sample Description of the Internet Survey

The sample of the Internet survey contained 454 participants; 55.1% (250/454) were male, 45.8% (208/454) between 21-30 years, and about 32.2% (146/454) were still at a university. There were 50.2% (228/454) that used the Internet often or very often to search medical information. There were 4.2% (19/454) participants that reported working in the medical sector. In total, the link of the survey was accessed 995 times, implying a completion rate of 45.5% (454/995) among those who had accessed the site. Slightly more of the 454 participants were assigned (51.1%, n=232) to the high-quality website. Analysis of the participants' Internet protocol (IP) addresses showed that all accessed the survey from a German Internet connection. The IP address is a unique number assigned of the computer used for the survey. A complete sample description is shown in Table 3.

No statistically significant differences could be found between male and female, age groups, Internet usage for health information, and educational levels. Working in the medical sector was negatively related to the ability to distinguish the quality of the website, but due to the small sample size, no further investigation can be done on this point.

**Table 3.** Detailed sample description of the Internet survey.

| Participants | | Total | | Exposure to high-quality page | | Exposure to low-quality page | |
|---|---|---|---|---|---|---|---|
| | | n | % | n | % | n | % |
| Total number, N | | 454 | 100 | 232 | 100 | 222 | 100 |
| **Gender** | | | | | | | |
| | Male | 250 | 55.1 | 111 | 47.8 | 72 | 32.4 |
| | Female | 171 | 37.7 | 99 | 42.7 | 139 | 62.6 |
| | Missing | 33 | 7.3 | 22 | 9.5 | 11 | 5.0 |
| **Age group, years** | | | | | | | |
| | 15-20 | 110 | 24.2 | 62 | 26.7 | 48 | 21.6 |
| | 21-30 | 208 | 45.8 | 90 | 38.8 | 118 | 53.2 |
| | 31-40 | 36 | 8.0 | 22 | 9.5 | 14 | 6.3 |
| | 41-50 | 36 | 8.0 | 15 | 6.5 | 21 | 9.5 |
| | 51-64 | 29 | 6.4 | 20 | 8.6 | 9 | 4.1 |
| | Missing | 35 | 7.7 | 23 | 9.9 | 12 | 5.4 |
| **Education** | | | | | | | |
| | No school degree | 1 | 0.2 | 1 | 0.4 | - | - |
| | In school education | 26 | 5.7 | 13 | 5.6 | 13 | 5.9 |
| | Some school degree | 59 | 13.0 | 37 | 15.9 | 22 | 9.9 |
| | High school degree | 82 | 18.1 | 44 | 19.0 | 38 | 17.1 |
| | Professional school degree | 4 | 0.9 | 2 | 0.9 | 2 | 0.9 |
| | In university education | 146 | 32.2 | 70 | 30.2 | 76 | 34.2 |
| | University degree | 135 | 29.7 | 64 | 27.6 | 71 | 32.0 |
| | Missing | 1 | 0.2 | 1 | 0.4 | - | - |
| **Working in the medical sector** | | | | | | | |
| | Yes | 19 | 4.2 | 6 | 2.6 | 13 | 5.9 |
| | No | 423 | 93.2 | 217 | 93.5 | 206 | 92.8 |
| | Missing | 12 | 2.6 | 9 | 3.9 | 3 | 1.4 |
| **Internet use for medical information** | | | | | | | |
| | Not at all | 13 | 2.9 | 7 | 3.0 | 6 | 2.7 |
| | 1 Little | 39 | 8.6 | 24 | 10.3 | 15 | 6.8 |
| | 2 | 86 | 18.9 | 44 | 19.0 | 42 | 18.9 |
| | 3 | 86 | 18.9 | 45 | 19.4 | 41 | 18.5 |
| | 4 | 66 | 14.5 | 33 | 14.2 | 33 | 14.9 |
| | 5 | 83 | 18.3 | 41 | 17.7 | 42 | 18.9 |
| | 6 | 41 | 9.0 | 20 | 8.6 | 21 | 9.5 |
| | 7 Very often | 38 | 8.4 | 16 | 6.9 | 22 | 9.9 |
| | Missing | 2 | 0.4 | 2 | 0.9 | - | - |

## Scale Reliability and Principal Component Analysis

By means of the principal component analysis, the different dimensions were tested and the number of items reduced. Out of the 45 items of the scale, four primary factors were identified accounting in total for 65% of overall variance, and following the analysis of the items' factor loadings and contexts, two factors were recognized as content-specific and the other two as website surrounding-specific factors. The 32 items, which are not part of the final scale, were excluded from further analysis as these displayed high cross-loadings, very low loadings, or no loadings on any factors. Factor 1 accounted for 32.37% (eigenvalue 4.275) of the variance, Factor 2 for 7.96% (eigenvalue 1.035), Factor 3 for 13.37% (eigenvalue 1.738), and Factor 4 for 10.83% (eigenvalue 1.408). The newly grouped items are shown in Table 4.

**Table 4.** Results of the principal component analysis.

| | Factors | | | |
|---|---|---|---|---|
| | Content-specific[a] | | Surrounding-specific[a] | |
| | 1 | 2 | 3 | 4 |
| The content convinced me. | .835 | | | |
| The website appears to be trustworthy. | .770 | | | |
| The website provides good information. | .758 | | | |
| The author seems to be knowledgeable due to the academic title. | .737 | | | |
| I learned something reading the content. | .688 | | | |
| The text is too long. | | .854 | | |
| The sentences have a difficult structure. | | .644 | | |
| Advertisements distracted me. | | | .796 | |
| The website contains dispensable links. | | | .732 | |
| Nothing distracts from the content. | | | .706 | |
| The website has a blurry layout. | | | .672 | |
| In general advertisement pop-ups help to add meaningful information. | | | | .853 |
| In general moving advertisement help to draw attention on the content. | | | | .726 |
| Rotation method, Varimax with Kaiser Normalization[b] | | | | |

[a] Extraction method, principal component analysis

[b] Rotated component matrix; Rotation converged in 5 iterations.

## Factor Labels

Factor 1 was labeled “Trustworthiness” and contained five items on the website being perceived as convincing, trustworthy, and informative (Cronbach alpha=.839). Factor 2 is “Textual deficits” and unites two items on sentence length and complexity (Cronbach alpha=.761). Factor 3, we called “Interference”; it binds items on irritation by advertisements, links, and layout (Cronbach alpha=.592). Finally, Factor 4, “Advertisements”, is on distraction or usefulness of advertisements (Cronbach alpha=.532).

The Promax rotation for four factors showed that there were no correlations higher than the threshold of .32. Following Tabachnick and Fidell [51], we continued with an orthogonal rotation. The results of the oblique rotation are shown in Table 5.

**Table 5.** Factor correlations of the principal component analysis.

| Factors[a] | 1 | 2 | 3 |
|---|---|---|---|
| 2 | .256 | | |
| 3 | -.157 | -.218 | |
| 4 | .052 | .067 | .198 |

[a] Rotation Method, Promax with Kaiser Normalization.

## Differences Between the Conditions

The analysis showed significant differences between the high- and the low-quality websites with regard to the perception of three of the four dimensions, all at a $P<.001$ significance level. Participants who had seen the high-quality website rated it higher on trustworthiness and interference, but lower on textual deficits. Regarding the fourth component, advertisements, both groups rated them as disturbing the users’ focus on content. The difference and the *t*-value were negative, but not significant ($P=.423$). The comparisons can be seen in Table 6. Taken together, these results show that the participants were able to distinguish the quality of medical information on the Internet with regard to trustworthiness and interference, whereas the low-quality site received better ratings on textual deficits. The perception of disturbing advertisements was unrelated to both exposures.

**Table 6.** Statistical differences between the two exposures.

| Components | M[a] | SD | df[b] | $t$ [c] | Significance |
|---|---|---|---|---|---|
| Trustworthiness (Factor 1) | .778 | .112 | 452 | 6.970 | <.001 |
| Interference (Factor 3) | .821 | .134 | 452 | 6.132 | <.001 |
| Textual deficits (Factor 2) | -.595 | .122 | 452 | -4.905 | <.001 |
| Advertisements[d] (Factor 4) | -.107 | .134 | 452 | -.802 | .423 |

[a] M=Mean

[b] df=degrees of freedom

[c] *t*=Student’s t distribution

[d] Equal variances not assumed for this item

### Impact of the Website Quality on the Outcome Measure

The reliability statistics for the four-item outcome measure (see Textbox 1) showed a Cronbach alpha=.853. To find out whether the participants would act differently depending on the quality of the website, an independent sample *t* test was conducted to evaluate the relationship of the outcome measure (Textbox 1) and the content quality. The Levene’s test showed that equal variances could not be assumed. The *t* test showed significant results $t_{446.806}$=5.519, *P*<.001. Participants rated the high-quality website (mean 4.46, SD 1.37) in the outcome measure to be better than the website with low-quality content (mean 3.73, SD 1.46). The 95% confidence interval for the difference in means was ranging from 0.47 to 0.99.

## *Discussion*

### Principal Findings

This research is based on the experience of average Internet users and quantitative testing of the designed scale. Therefore, it was possible to design a novel measure that covers, on the one hand, similar aspects as the DISCERN scale, but provides, on the other hand, important additional Internet-specific items. The items of the widely used DISCERN measure are divided into two sections that focus on the concepts of quality and credibility of the given information [33]. The items of the presented measure share the importance of constructs measured in the DISCERN, but work differently. In contrast to the existing measure, the items’ structure in the proposed measure is more general and easier for laypersons to assess. It is relevant by taking the particular behavior of Internet information users into consideration. Written information on the Internet can be described as more viral and superficial compared to information found on other sources of mass media, in particular considering the surrounding-specific factors’ interference and advertisements [35,49,50]. The proposed novel measure focuses exclusively on the impressions Internet-users get from the consulted websites. This notion is to date not covered by measures such as DISCERN, but crucial for assessing the credibility impressions of consumers. This proposed measure was developed, therefore, to reflect the behaviors and decisions of individuals searching for health information. In contrast to previously mentioned measures, we did not use samples of individuals with special expertise or professional medical background, but focused exclusively on average Internet-users. Taking together these characteristics, the proposed measure can be combined with existing measures [36-38] on the credibility assessment of health information on the Internet to explore, in a next step, the user perception of the provided health information.

The sufficient level of scale reliability and the properties of this measure suggest that this measure allows examining the view of health information seekers on the provided information. The experimental design showed that the ratings developed for the scale differentiate between a high- and a low-quality website. This makes this measure a useful tool for examining patients’ Internet searches. The measure was not designed based on specific websites, but on the search procedures of the participants of the observational study. Moreover, it is not condition-specific and can be administered to all medical information websites on the Internet. These characteristics allow administering the tool relatively easily in either Internet- or in paper and pencil-based research studies. It can thus be an easy to use measuring tool, which can be incorporated alongside other measures. Useful apps can be found in the eHealth area and for website testing in health campaigns.

Typical for the experimental research layout, several aspects worked differently from what we expected. Between the two experimental groups, the results showed that participants who were exposed to the high-quality website rated its credibility in this measure higher on the factors trustworthiness and interference, but lower on textual deficits. The unexpected direction of the difference could be due to the different styles of the sites. While the high-quality site had long explanatory text parts, the low-quality site had only simple information. Moreover, unexpected results were found on Factor 4 grouping the advertisement items. The nonsignificant results for the correlation of the experimental conditions seem to be reflected within the specific item wording. In contrast to all other items in the final measures, these items could have suggested a more general answer by the participants, which was not limited to the context of the website they had seen. Participants answered this item based on their general attitude and opinion, and consequently, the answers were not affected by the website they had seen. This is reflected by the nonsignificant results of this factor.

Most of the results regarding the rating of the different quality of websites matched with the previous assumption of the research group. For this case, the measure seems to provide a sufficient rating tool able to produce judgments consistent with experts' categorizations. Although the testing in this study was done on sleeping disorder websites, other conditions can be included. As the measure is by its content not bound to a specific disease or medical condition, it can be widely used. With respect to the growing usage of Internet apps and Internet information by health professionals and laypersons, the measurement catalogue is still very limited when it comes to the combination of content quality and medical information.

### Limitations and Further Research

Initiating a research project with a student sample caused some difficulties overcome by using the snowball system in order to include participants from outside the university. Still, the average age of the sample is rather young and, therefore, does not represent the society of Internet users. It should also be mentioned that health information searches on the Internet are linked to such sociodemographic characteristics as age, gender, and health status [9,10,35]. The presented measure can only be applied to a specific website, but it does not help to understand other relevant determinants such as the result presentation by the search engines. Moreover, the final measure is the result of a statistical analysis, which showed significant effects, but lacks further testing as a composite measure, and, therefore, should be interpreted with caution. This research focused exclusively on one medical condition and did not test the measure with other conditions, which would allow proving the consistency of the measure across different medical subjects. With regard to this aspect, it is unclear how far the measure produces reliable results when considering such controversial medical topics as vaccination or cancer treatment.

Further research with another independent sample will allow confirming the factor structure of the scale. Moreover, it would be possible to provide solutions to some of the limitations and to improve the measure by defining cut-point values as estimators for high- or low-quality content of websites. The measure would in this way offer the possibility of addressing health information users on the Internet who struggle with identifying quality websites. It would also be practical to continue examining this measure in comparison with the health literacy levels of participants to see whether predictors can be found there. So far, the results showed that (formal) knowledge did not show any differences in the research population.

### Conclusions

This measure provides a practical tool, which will show its relevance for research on health information on the Internet. In contrast to previous attempts, this measure is designed for the Internet-setting of this information channel and the particular users' behavior. The inclusion of the laypersons' experience into the measurement development process might be seen as unusual, but crucially, this brings the consumers' perspective into academic research. Therefore, the initially mentioned concern of public health authorities on the quality of health information provided on the Internet [10] can be answered, and the result of this research offers a tool for assessing user perception of content quality. Unlike other information, the impact and the consequences of health information have the potential to be severe. Across gender, age group, and educational level, this measure provides a clear answer on the abilities of participants to estimate the quality of medical information on the Internet. Website testing can be enriched by a credibility criterion based on the ratings of participants. As the amount of medical information on the Internet increases and patients are increasingly empowered to decide on relevant health matters, the research link between general quality assessment and Internet health information becomes relevant. The skill to critically consume health information is important to fully make use of the opportunities and health benefits which eHealth tools offer. From a scientific point of view, the disparities, which can be seen in health literacy levels, will probably be the same when it comes to medical information usage on the Internet. Therefore, understanding how participants perceive medical information on the Internet is a first step to identifying needs and addressing them properly. A measure is ready to be used for the assessment of patients' perception of credibility of eHealth contents.

## Acknowledgments

The authors wish to thank Professor Dr Hans-Bernd Brosius, the Center of Advanced Studies, and the undergraduate students from the Ludwigs-Maximilians Universität Munich, Germany for their invitation and the possibility to give a course and their support of this study.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

The seven dimensions with the original items of the measure based on the observational study (compare Table 1 in the article) and the original items of the outcome measure (compare Textbox 1 in the article).

[PDF File (Adobe PDF File), 8KB - ijmr_v4i1e8_app1.pdf ]

## References

1. 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]
2. Baker L, Wagner TH, Singer S, Bundorf MK. Use of the Internet and e-mail for health care information: Results from a national survey. JAMA 2003 May 14;289(18):2400-2406. [doi: 10.1001/jama.289.18.2400] [Medline: 12746364]
3. Fox S, Duggan M. Health online. Washington, DC: Pew Internet & American Life Project; 2013. URL: http://pewinternet.org/Reports/2013/Health-online.aspx [WebCite Cache ID 6WoCtKkvl]
4. Hesse BW, Moser RP, Rutten LJ. Surveys of physicians and electronic health information. N Engl J Med 2010 Mar 4;362(9):859-860. [doi: 10.1056/NEJMc0909595] [Medline: 20200398]
5. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: The impact of the Internet and its implications for health care providers: Findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
6. Bader JL, Theofanos MF. Searching for cancer information on the internet: Analyzing natural language search queries. J Med Internet Res 2003 Dec 11;5(4):e31 [FREE Full text] [doi: 10.2196/jmir.5.4.e31] [Medline: 14713659]
7. Eysenbach G, Köhler C. What is the prevalence of health related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. In: http://europepmc.org/abstract/MED/14728167. 2003 Presented at: AMIA Annu Symp Proc; 2003; USA p. 225-229.
8. Fox S, Fallows D. Internet health resources. Washington, DC: Pew Internet & American Life Project; 2003. URL: http://www.pewinternet.org/2003/07/16/internet-health-resources/ [accessed 2006-08-08] [WebCite Cache ID 6Wx25xgj8]
9. Kummervold PE, Chronaki CE, Lausen B, Prokosch HU, Rasmussen J, Santana S, et al. eHealth trends in Europe 2005-2007: A population-based survey. J Med Internet Res 2008;10(4):e42 [FREE Full text] [doi: 10.2196/jmir.1023] [Medline: 19017584]
10. Broom A. Virtually he@lthy: The impact of internet use on disease experience and the doctor-patient relationship. Qual Health Res 2005 Mar;15(3):325-345. [doi: 10.1177/1049732304272916] [Medline: 15761103]
11. Reavley NJ, Jorm AF. The quality of mental disorder information websites: A review. Patient Educ Couns 2011 Nov;85(2):e16-e25. [doi: 10.1016/j.pec.2010.10.015] [Medline: 21087837]
12. Lorence DP, Greenberg L. The zeitgeist of online health search. Implications for a consumer-centric health system. J Gen Intern Med 2006 Feb;21(2):134-139 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0295.x] [Medline: 16336621]
13. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet. JAMA 2001 May 23;285(20):2612. [doi: 10.1001/jama.285.20.2612]
14. Nutbeam D. Health literacy as a public health goal: A challenge for contemporary health education and communication strategies into the 21st century. Health Promot Int 2000;15(3):259-267 [FREE Full text] [doi: 10.1093/heapro/15.3.259]
15. Becker PM. Insomnia: Prevalence, impact, pathogenesis, differential diagnosis, and evaluation. Psychiatr Clin North Am 2006 Dec;29(4):855-870; abstract vii. [doi: 10.1016/j.psc.2006.08.001] [Medline: 17118272]
16. Attarian HP, Perlis ML. Defining insomnia. In: Attarian HP, Schuman CS. editors. Clinical Handbook of Insomnia. 2nd ed. New York: Springer; 2010:3-12.
17. Grewal R, Doghramji K. Epistemology of insomnia. In: Clinical Handbook of Insomnia. 2nd ed. New York: Springer; 2010:13-22.
18. Lichstein K, Durrence H, Taylor D, Bush A, Riedel B. Quantitative criteria for insomnia. Behaviour Research and Therapy 2003 Apr;41(4):427-445. [doi: 10.1016/S0005-7967(02)00023-2]
19. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web. JAMA 2002 May 22;287(20):2691. [doi: 10.1001/jama.287.20.2691]
20. Stvilia B, Mon L, Yi YJ. A model for online consumer health information quality. J. Am. Soc. Inf. Sci 2009 Sep;60(9):1781-1791. [doi: 10.1002/asi.21115]
21. Bernstam EV, Walji MF, Sagaram S, Sagaram D, Johnson CW, Meric-Bernstam F. Commonly cited website quality criteria are not effective at identifying inaccurate online information about breast cancer. Cancer 2008 Mar 15;112(6):1206-1213 [FREE Full text] [doi: 10.1002/cncr.23308] [Medline: 18266210]
22. Gustafson DH, Wyatt JC. Evaluation of ehealth systems and services. BMJ 2004 May 15;328(7449):1150 [FREE Full text] [doi: 10.1136/bmj.328.7449.1150] [Medline: 15142895]
23. Purcell GP, Wilson P, Delamothe T. The quality of health information on the internet. Br Med J 2002;324(7337):557-558. [doi: 10.1136/bmj.324.7337.557]
24. Wilson P. How to find the good and avoid the bad or ugly: A short guide to tools for rating quality of health information on the internet. Br Med J 2002 Mar;324(7337):598-602. [doi: 10.1136/bmj.324.7337.598]
25. Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the Internet. JAMA 2001 May 23;285(20):2612. [doi: 10.1001/jama.285.20.2612]
26. Hardey M. "E-health": The internet and the transformation of patients into consumers and producers of health knowledge. Info., Comm. & Soc 2001 Oct;4(3):388-405. [doi: 10.1080/713768551]
27. Eysenbach G. Consumer health informatics. Br Med J 2000;320(7251):1713-1716. [doi: 10.1136/bmj.320.7251.1713]

28. Jadad AR, Gagliardi A. Rating health information on the Internet. JAMA 1998 Feb 25;279(8):611. [doi: 10.1001/jama.279.8.611]
29. Maloney S, Ilic D, Green S. Accessibility, nature and quality of health information on the Internet: A survey on osteoarthritis. Rheumatology (Oxford) 2005 Mar;44(3):382-385 [FREE Full text] [doi: 10.1093/rheumatology/keh498] [Medline: 15572390]
30. Ilic D, Risbridger G, Green S. Searching the Internet for information on prostate cancer screening: An assessment of quality. Urology 2004 Jul;64(1):112-116. [doi: 10.1016/j.urology.2004.02.027] [Medline: 15245946]
31. Kunst H, Groot D, Latthe PM, Latthe M, Khan KS. Accuracy of information on apparently credible websites: Survey of five common health topics. Br Med J 2002 Mar;324(7337):581-582. [doi: 10.1136/bmj.324.7337.581]
32. Khazaal Y, Chatton A, Cochand S, Coquard O, Fernandez S, Khan R, et al. Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. Patient Educ Couns 2009 Oct;77(1):33-37. [doi: 10.1016/j.pec.2009.02.016] [Medline: 19372023]
33. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. Journal of Epidemiology & Community Health 1999 Feb 01;53(2):105-111. [doi: 10.1136/jech.53.2.105]
34. 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,in-depth interviews. Br Med J 2002 Mar;324:573-577. [doi: 10.1136/bmj.324.7337.573]
35. Rains SA, Karmikel CD. Health information-seeking and perceptions of website credibility: Examining Web-use orientation, message characteristics, and structural features of websites. Computers in Human Behavior 2009 Mar;25(2):544-553. [doi: 10.1016/j.chb.2008.11.005]
36. Eysenbach G, Yihune G, Lampe K, Cross P, Brickley D. MedCERTAIN: Quality management, certification and rating of health information on the Net. In: Proc AMIA Symp. 2000 Presented at: AMIA Annu Symp Proc; 2000; USA p. 230-234 URL: http://europepmc.org/abstract/MED/11079879
37. Eysenbach G, Thomson M. Stud Health Technol Inform. 2007. The FA4CT algorithm: A new model and tool for consumers to assess and filter health information on the internet URL: http://ebooks.iospress.nl/volume/medinfo-2007 [accessed 2015-03-04] [WebCite Cache ID 6WoFVGsAK]
38. Eysenbach G, Köhler C, Yihune G, Lampe K, Cross P, Brickley D. A metadata vocabulary for self- and third-party labeling of health web-sites: Health information disclosure, description and evaluation language (HIDDEL). In: Proc AMIA Symp. 2001 Presented at: AMIA Annu Symp Proc; 2001; USA p. 169-173 URL: http://europepmc.org/abstract/MED/11825174
39. DeVellis RF. Scale development: Theory and applications. Thousand Oaks, Calif: Sage; 2003.
40. DeWalt KM, DeWalt BR. Participant observation: A guide for fieldworkers. USA: Walnut Creek, CA; Alta Mira Press; 2002.
41. Schensul SL, LeCompte M, Schensul S. The ethnographer's toolkit. Walnut Creek, Calif: Altamira; 1999.
42. Bogdan RB, Biklen S. Qualitative research for education: An introduction to theories and methods. Boston, Mass: Pearson A & B; 2007.
43. Stiftung W. Test. 2009. Gesundheitsportale im Internet URL: https://www.test.de/Gesundheitsportale-Die-besten-Infos-im-Netz-1780855-0/ [accessed 2015-03-05] [WebCite Cache ID 6WoG1drL1]
44. Psychometrica Lesbarkeitsindex (LIX). Available at: URL: http://www.psychometrica.de/lix.html [accessed 2015-02-25] [WebCite Cache ID 6WbmYsGrO]
45. Thurstone LL. Multiple factor analysis. Chicago, IL: Chicago University Press; Dec 1949.
46. Gerbing DW, Hamilton JG. Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. Structural Equation Modeling: A Multidisciplinary Journal 1996 Jan;3(1):62-72. [doi: 10.1080/10705519609540030]
47. Comrey AL, Lee HB. A first course in factor analysis. Hillsdale, N.J: Erlbaum; 1992.
48. IBM SPSS Statistics computer software. Version 21. Armonk: IBM; 2013. URL: http://www-01.ibm.com/support/docview.wss?uid=swg24032236 [WebCite Cache ID 6WwmVpvlr]
49. O'Keeffe J, Willinsky J, Maggio L. Public access and use of health research: An exploratory study of the National Institutes of Health (NIH) Public Access Policy using interviews and surveys of health personnel. J Med Internet Res 2011;13(4):e97 [FREE Full text] [doi: 10.2196/jmir.1827] [Medline: 22106169]
50. Graham L, Tse T, Keselman A. NCBI. 2006. Exploring user navigation during online health information seeking URL: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1839467/ [accessed 2015-03-05] [WebCite Cache ID 6WoGWJiDj]
51. Tabachnick BG, Fidell LS. Using multivariate statistics. Boston: Pearson/Allyn & Bacon; 2007.

## Abbreviations

**HTML:** HyperText Markup Language
**IP:** Internet protocol

*Edited by G Eysenbach; submitted 28.11.13; peer-reviewed by R Ownby, J Walker; comments to author 27.04.14; revised version received 27.05.14; accepted 12.01.15; published 30.03.15.*

*Please cite as:*
*Dubowicz A, Schulz PJ*
*Medical Information on the Internet: A Tool for Measuring Consumer Perception of Quality Aspects*
*Interact J Med Res 2015;4(1):e8*
*URL: http://www.i-jmr.org/2015/1/e8/*
*doi:10.2196/ijmr.3144*
*PMID:25835333*

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