

Characteristics of online and offline health information seekers and factors that discriminate between them

Shelia R. Cotten*, Sipi S. Gupta

Department of Sociology and Anthropology, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA

Abstract

Increasing number of individuals are using the internet to meet their health information needs; however, little is known about the characteristics of online health information seekers and whether they differ from individuals who search for health information from offline sources. Researchers must examine the primary characteristics of online and offline health information seekers in order to better recognize their needs, highlight improvements that may be made in the arena of internet health information quality and availability, and understand factors that discriminate between those who seek online vs. offline health information. This study examines factors that differentiate between online and offline health information seekers in the United States. Data for this study are from a subsample ($n = 385$) of individuals from the 2000 General Social Survey. The subsample includes those respondents who were asked Internet and health seeking module questions. Similar to prior research, results of this study show that the majority of both online and offline health information seekers report reliance upon health care professionals as a source of health information. This study is unique in that the results illustrate that there are several key factors (age, income, and education) that discriminate between US online and offline health information seekers; this suggests that general “digital divide” characteristics influence where health information is sought. In addition to traditional digital divide factors, those who are healthier and happier are less likely to look exclusively offline for health information. Implications of these findings are discussed in terms of the digital divide and the patient–provider relationship.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Health information seeking; Internet; Digital divide

Introduction

With the fast-paced progression of the information revolution and greater accessibility to the World Wide Web (WWW), individuals are increasingly turning to the Internet to satisfy a variety of informational, communication, and entertainment needs. Its speedy expansion has generated “an information revolution of unprecedented magnitude” (Jadad & Gagliardi, 1998, p. 611). Individuals utilize this powerful tool to keep in touch with and expand social support networks, keep updated with sporting and news events on an international level,

and search for various types of information to assist them in their daily lives. In addition, the Internet has been identified as a practical tool for connecting with many low-income, less educated, and minority individuals to disseminate health information (Zarcadoolas, Blanco, Boyer, & Pleasant, 2002; Brodie, Altman, Blendon, Benson, & Rosenbaum, 2000).

Prior research indicates that there is great potential to utilize the Internet to aid in disseminating health information to the public at large (Fox & Fallows, 2003; Cotten, 2001; Brodie et al., 2000). As this medium for conveying information continues to grow in accessibility and popularity, a more complete understanding is required of the extent to which online health information is penetrating people’s lives, the scope of the impact it will have on them, and implications of

*Corresponding author. Tel.: +410-455-6330; fax: +410-455-1154.

E-mail address: cotten@umbc.edu (S.R. Cotten).

seeking online health information. Above all, we must discover the primary characteristics of online and offline health information seekers to better recognize their needs, highlight improvements that may be made in the arena of Internet health information quality and availability, and understand factors that discriminate between those who seek online vs. offline health information.¹

As a first step in this effort, the purpose of this study is to examine how individual characteristics, health status, and computer and Internet usage affect health information seeking among a sample of United States adults. While previous research has begun to investigate the characteristics of individuals in the US who use the Internet (see for instance, the Pew Internet & American Life Studies by Kommers & Rainie, 2002; Rainie & Packel, 2001; Spooner & Rainie, 2000), little is known about how computer and Internet usage, health status, and sociodemographic characteristics affect health information seeking. These results will help us discern important differences between those who seek health information through the Internet compared to those who do not.

Health information seeking

Health information seeking is defined as the search for and receipt of messages that help “to reduce uncertainty regarding health status” and “construct a social and personal (cognitive) sense of health” (Tardy & Hale, 1998, p. 338). Traditionally, physicians have been the purveyors and integrators of health care information (Goldsmith, 2000), and thus they have held almost exclusive access to this information or expert knowledge (Giddens, 1991; Hardey, 1999). Given that patients have held less power in the health care arena, the doctor–patient relationship has been inherently unequal (Shackley & Ryan, 1994).

Despite a willingness on their part to obtain health-related information, patients are not likely to display information seeking behavior in the company of their doctors (Beisecker, 1988). This may be a function of either the hierarchical relationship between patients and providers and/or the declining length of appointment times. Fox and Rainie (2000) suggest that being informed health care consumers will become even more important as health management organizations become less lenient and the burden of responsibility for patient knowledge and choices falls increasingly upon the patient.

¹ Much of the literature reviewed in this manuscript pertains to findings from US based studies and may not apply to other populations.

Apart from health care providers, traditional modes of health information seeking include the use of local experts and the mass media. Specifically, individuals seek “local experts” (i.e., individuals who have tangible experience in the health care profession or who themselves once experienced the medical condition) for advice on health-related issues. These lay individuals often appear approachable and amicable and are integrated into their local communities. Thus people often feel more comfortable seeking health information from them than from their health care providers (Tardy & Hale, 1998).

In addition to local experts, many individuals rely upon various forms of mass media for health information (Brodie, Kjellson, Hoff, & Parker, 1999). Mass media sources include magazines, newspapers, other printed publications, television, radio, street signs/billboards, and in recent years, the Internet. Analysis of three national surveys reveals that mass media was cited as the predominant source of health news for the majority of the respondents in these studies (Brodie et al., 1999). Other, more recent, studies also report similar findings. For instance, African–American youth reported acquisition of information on the health effects of smoking through television viewing (Kurtz, Kurtz, Johnson, & Cooper, 2001) and television news and entertainment shows are often cited by consumers as sources of health information (Brodie et al., 2001). Montagne (2001) reports that the recent increase in Prozac use is the result of a heightened media-generated drug information and awareness campaign. Mass media plays a substantial role in defining health and illness, detailing products and services designed to assist individuals in negotiating their health and well-being, and providing models of others with particular health concerns for consumers.

Health information seeking and the Internet

Although television and other media sources play pivotal roles in the dissemination of health information, the Internet may be subsuming much of this function. According to Goldsmith (2000, p. 148), “the Internet exploded during the late 1990s into a powerful new social institution.” It is now a heavily relied upon source of reference material for the public that transcends existing geographical and regulatory boundaries, and where distinctions between professions and expertise are blurred (Hardey, 1999). The Internet has rapidly expanded to address the demand for medical information on health-related topics (McLeod, 1998; Patrick, 2000), and health information is widely prevalent and often sought on the Internet (Fox & Fallows, 2003; Cotten, 2001; Harris Interactive, 2001; Suarez-Amazon, Kendall, & Dorgan, 2001; Elliott & Elliott, 2000) by

both consumers and providers of health services (Jadad & Gagliardi, 1998).

A variety of mediums exist through which individuals can access health information online. These include websites, listservs, online support groups, chat rooms, instant messaging, and email. More than 70,000 websites contain health information (Grandinetti, 2000), and the number of health websites is rapidly increasing (Nielsen//NetRatings, 2002). In addition to website utilization and Internet consultation with medical professionals, online support communities have been identified as one of the primary methods of online health information seeking for both consumers and members of their social networks (Kummervold et al., 2002; Cline & Haynes, 2001; Preece, 2000; Cotten, 2001). The importance of such support networks has been documented in studies of Alzheimers patients, individuals with knee injuries and breast cancer, and caregivers of cancer patients (to name just a few) (Preece, 2000). The understanding that others are in similar situations and the empathy this affords has been instrumental in the strengthening of these communities (Preece, 2000).

Although a multitude of online health information exists, there is little consistency in terms of how many people actually use this information, their expectations of it, and the implications of this usage. Estimates of individuals utilizing the Internet for health information vary depending upon the sampling design and methodology used. However, the most recent research from the Pew Internet Project suggests that the figures have increased by almost 20% in just a year, with 80% of adult Internet users in the US (93 million) utilizing the Internet to obtain health-related information (Fox & Fallows, 2003). Regardless of the source, it is evident that those seeking health information via the Internet are part of a growing sector, and one that is expected to increase in coming years (Cotten, 2001; Taylor, Alman, & Manchester, 2001; Fattah, 2000).

Advantages and disadvantages of online health information seeking

In many ways, the Internet is an optimal way to disseminate health information. It affords users immediate access to an incredible amount of health-related information that is directed towards both health care professionals and the general public (Bernhardt, Lariscy, Parrott, Silk, & Felter, 2002; McLeod, 1998). It affords individuals privacy, immediacy, convenience, anonymity, a wide variety of information, and a variety of perspectives on the same topic (Skinner, Biscope, & Poland, 2003; Fox & Rainie, 2000; Bischoff & Kelley, 1999). The cloak of confidentiality afforded by the anonymous nature of the Internet is advantageous in

that it allows users to ask awkward, sensitive, or detailed questions without the risk of facing judgment, scrutiny, or stigma, and to do so at their convenience. Additional impetuses for the growth of online health information seeking include the popularization of the participative health care model, the difficulty any one physician would face in keeping abreast of the sizeable and ever-increasing amount of available health information, limited office-visit time due to cost-containment measures, increased attention to prevention and self-care, the greater health care needs of an aging population, and a growing interest in alternative medicine (Cotten, 2001; Eng et al., 1998; Gallagher, 1999). The immediacy of information access, the accessibility at any time of the day or night, the potential continual updating of information, and the wider range of information available positively distinguish the Internet from other forms of information services (e.g., the newspaper, radio, and television).

Although there are many advantages to online health information seeking, disadvantages have also been identified. Internet and website accessibility issues continue to plague many with health impairments (Davis, 2002). In addition, privacy and anonymity concerns are key issues for individuals seeking online health information (Bernhardt et al., 2002). “Health seekers” are opposed to Internet tracking, fearful that it may reveal their searches and that their insurance companies or employers may become privy to this information. They also object to the notion of medical records being posted online, even if a secure server is being used. Moreover, only 21% of health seekers have disclosed their email address when visiting a health-related website, and only 17% have provided personal information at these sites (Fox & Rainie, 2000). Others have argued that the Internet’s information overload, general disorganization, use of overly technical language, lack of user friendliness, and constant changeability can increase searching difficulties (Jadad & Gagliardi, 1998; Cline & Haynes, 2001).

Credibility of health information on the Internet is also a valid concern (Bernhardt et al., 2002), with 83% of health seekers in one study emphasizing the importance of gathering information from multiple sites (Fox & Rainie, 2000). Due to a lack of sufficient online health-content regulation, consumers are at special risk of becoming victims of “official looking” web pages that lack peer review (Cline & Haynes, 2001). Research by Impicciatore and Pandolfini (1997) showed that only nine out of 41 surveyed websites provided correct instructions regarding temperature-taking in children. Online support groups have also been implicated as culprits in the dissemination of misleading information, because much of the guidance offered by support group members is based on personal experience and often lacks

the critical perspective developed by health professionals, who are required to discriminate among resources to determine information quality (Cline & Haynes, 2001). These findings underscore the need for consumers to validate online health information.

Consumer characteristics

Although most individuals perceive doctors and other health care professionals as a main source of health information (Pennbridge, Moya, & Rodrigues, 1999), mass media plays an important role in the education of individuals on health-related topics. As some researchers note, however, the usage and impact of mass media can vary depending on consumer characteristics (Kakai, Maskarinec, Shumay, Tatsumura, & Tasaki, 2003; Tardy & Hale, 1998). Brodie et al. (1999) report that the majority of whites, African Americans, and Latinos rely upon the media as a source for health related information. However, economically disadvantaged individuals may not be able to fully utilize mass media resources, due to lower reading levels, limited access to new information sources, and an external locus of control with regard to their own health problems (Tardy & Hale, 1998; Pirisi, 2000). Those who more frequently engage the services of mass media are better able to discriminate between useful and non-useful information, have more acceptance and retention of information, and possess better communication skills (Tardy & Hale, 1998). However, individuals are more likely to give greater credibility to information obtained from a computer than to that obtained from other media sources (Cline & Haynes, 2001).

For many Americans under the age of 60, computer and Internet usage have become standard elements of everyday life (Brodie et al., 2000). However, a large sector of the population remains detached from this experience. Disparities in computer access and usage have been found within income, education, age, and ethnic groups (NTIA, 2000; Bolt & Crawford, 2000). This gap, commonly referred to as the “digital divide,” especially affects lower-income African Americans. When controlling for income, however, the racial gap declines somewhat (Brodie et al., 2000). From initial studies it appears that health information seeking behavior is appealing to individuals regardless of race or income status, especially as compared to other online search patterns (Fox & Rainie, 2000). However, African Americans appear to rely on the Internet’s capabilities as an informational tool more so than whites (45% vs. 35%) (Spoonster & Rainie, 2000).

Although few studies have examined the characteristics of online health information seekers, the few that have been done show that the more experienced an Internet user is, the more likely s/he is to search for

health information online (Fox & Fallows, 2003; Fox & Rainie, 2000). On average, health information sought online by Internet users takes less than 6 min to obtain, and individuals often report success in uncovering sought-after information (Eysenbach & Kohler, 2002). Most of these online health seekers search for specific information on specific conditions (93%) and many report gathering information prior to visiting their health care provider (55%) (Fox & Rainie, 2002). Approximately 30% of respondents in another study reported seeking information related to health care providers, while 19% reported seeking information on sexual health issues (Brodie et al., 2000). Fox and Rainie (2002) note that 33% of online health seekers in their study have sought information about a sensitive health topic that they would find hard to discuss with others, while Kummervold et al. (2002) find that almost half of their respondents discuss personal problems online that they do not discuss when they are face-to-face with others.

Research shows that less healthy people are more apt to use the Internet to find information relating to physical and mental health issues, compared to those who report better health status (Fox & Rainie, 2000). This finding supports the health needs model developed by Pandey, Hart, and Tiwary (2002). In addition, online health seekers are highly likely to have health insurance (91%) and over half report seeking health information on behalf of someone else (54%) (Fox & Rainie, 2000). Seventy percent report seeking specific information, often in response to a recent doctor’s visit by themselves or others. The fact that the majority of health seekers are likely to look for information on behalf of someone else, usually a family member or other loved one, may be due in part to the fact that 17% of them are Baby Boomers, with responsibilities for both parents and children (Fattah, 2000).

Studies also show that women are more likely to seek health information online than are men (Fox & Fallows, 2003; Hern, Weitkamp, Hillard, Trigg, & Guard, 1998; Fox & Rainie, 2000). Men and women have other differences in their online health seeking patterns. Women are likely to conduct Internet searches focused on an illness or its symptoms and, as the most active health seekers, are more likely to register strong positive beliefs regarding the benefits of online health searches. Men are more likely than women to allow Internet information to affect their searches,² and they are less concerned than are women about the credibility of health information found online. Men’s searches tend to

²See Bastardi and Shafir (1998), “On the Pursuit and Misuse of Useless Information” for a fascinating study regarding the weight non-instrumental information can play in affecting major decisions when it is actively pursued (as in an Internet search).

focus on the prognosis and treatment of a disease, and they are more inclined than women to use their newfound information when asking follow-up questions of their providers (Fox & Rainie, 2000).

While research in the area of health information seeking and the Internet is increasing, many questions remain to be answered. Many of the studies in this area, particularly those by the Pew Internet & American Life Project, are largely descriptive in nature. In addition, few studies have examined factors that discriminate between where individuals seek health information. Further multivariate level research is needed that investigates the interrelationships among Internet usage, sociodemographic characteristics, health status, and health information seeking. It is also important to control for the number of sources utilized, as it may be that those who utilize the Internet are primarily individuals who look for health information from many sources.³ It is not enough to simply note associations between health information seeking and Internet usage; additional research is needed that controls for key factors that have been related to health information seeking in prior research. To our knowledge, this is the first empirical study that combines these factors to examine how they relate to where individuals seek health information. Based on prior research, we hypothesize that characteristics of individuals, such as age, income, race, sex, Internet usage, and health status will distinguish those who seek health information online from offline health information seekers. Specifically, we hypothesize that those who are older, those with lower incomes, minorities, and males will be less likely to be in the group that uses the Internet for health information seeking. Individuals who report higher levels of Internet use should be more likely to be in the online health information group than those reporting less Internet use. We also hypothesize that those with worse health status would be more likely to use the Internet for health information.

Methods

Data and sample

Data for this study are derived from the 2000 General Social Survey (GSS), which was carried out by NORC. The GSS covers a wide range of variables, and the 2000 GSS included a module of questions concerning computer and Internet usage.

Sampling consisted of a national full probability sample of 2817 non-institutionalized adults in the

United States. Surveys were administered via 90-min in-person interviews, with a response rate of 70%. Because of the high response rate, we do not view non-response error as an area of significant concern. However, due to the administration methodology of the GSS, not all respondents were administered all modules of the GSS; thus when examining questions from various modules of the GSS, the sample size may be significantly smaller than the total sample size reported above. The subsample of respondents for the current study consists of only those respondents who had sought health information in the past year and were asked questions from the Internet and health-related modules of the GSS ($n = 385$). Details of both the original sample and the subsample used for this study are included in Table 1 and discussed in the Results section.

Measures

Key measures included in this study consist of health media usage, sociodemographic factors, Internet and computer usage measures, and health and happiness status. As drafts of the core questions of the GSS are reviewed and commented upon by approximately 150 social scientists, the survey measures should be reliable and valid.

Health information seeking: The key outcome for this study was a measure that assessed whether individuals sought health information offline vs. online. This measure was created through the use of a series of questions that asked respondents where they had sought information regarding a health concern or a medical problem in the past year. These included questions that asked respondents whether they had sought health information from (1) newspaper articles, (2) general interest magazines, (3) health magazines, (4) doctors or nurses, (5) friends or relatives, (6) the television or radio, or (7) the World Wide Web. For each of these measures, response options included: (1) not at all, (2) one or two times, and (3) three or more times. Those respondents who reported only having sought health information through offline sources were categorized as *offline or non-Internet* health information seekers ($n = 158$). Individuals who reported seeking health information solely online or those who used the Internet in combination with other sources were classified as *online or Internet* health information seekers ($n = 170$).⁴

A new variable was created to assess the number of sources from which individuals sought health

³We thank one of the reviewers for suggesting the importance of controlling for the number of information sources utilized.

⁴Initially, the intent was to examine three groups of health information seekers: online, offline, and combination online/offline health information seekers. However, very few respondents reported using strictly online sources for health information. Thus, two groups are used in the current study.

information. The aforementioned seven sources were used to create this measure. If respondents reported yes to seeking health information from that source, they received a score of one for that source; if not, they received a zero. The scores were summed and the respondents could receive a score from zero to seven on this measure.

Sociodemographic factors: Independent variables such as sex, ethnicity, age, total family income, marital status, and education were examined, as they have traditionally been associated with Internet usage and digital divide issues (NTIA, 2000). Sex was measured as male and female. Racial categories included white, black, and other. For purposes of this study, race was recoded into white vs. other. Age was reported in years. Income was assessed as total family income and was measured by twelve ranges of categories, with a low score of less than \$1000/year and a high score of \$25,000 or more/year. Given the distribution of the income variable, it was recoded into a dummy variable with those making less than \$25,000 per year scored zero and those making \$25,000 and higher scored one. Marital status was assessed as married, widowed, divorced, separated, or never married; however it was recoded into a dummy variable comparing non-married and married. Education was measured by the highest year of school completed.

Internet and computer usage: Two variables measured whether respondents used computers at work and at home. The term “at work” was defined as “a workplace away from your home.” Response options for the two measures were (1) yes and (2) no. Internet usage was assessed through two measures. Respondents were asked whether they used the Internet for non-email related purposes (yes/no response options), and the amount of time spent using email was assessed through an open-ended question asking for hours per week.

Well-being: Self-reported health status and general happiness were used as indicators of well-being. The health status measure has been used extensively and has been shown to be highly predictive of physical impairment. Respondents were asked to describe their own health. Response categories were (1) excellent, (2) good, (3) fair, and (4) poor. A dummy variable was created with those scoring one having excellent or good health, and those scoring zero having fair or poor health. Respondents also gave information about their general happiness. Response options for this question were (1) very happy, (2) pretty happy, and (3) not too happy. These categories were collapsed into happy (scored one) and unhappy (scored zero).

Analytical design

Univariate analysis is used to examine the frequency and distribution of study variables. Bivariate analyses

(χ^2 - and *t*-tests) are used to examine whether differences exist between the *online* and *offline* health information seekers in terms of the predictor variables.

Multivariate discriminant analysis is used to examine the hypothesized factors that discriminate between those who seek *online* vs. *offline* health information. Discriminant analysis uses multiple predictor variables to discriminate among individuals based upon a single dichotomous criterion variable. The single dichotomous criterion variable is the dummy variable (0 = offline health information seeking; 1 = online health information seeking) that gauges where individuals sought health information in the past year. In this analysis, the multiple predictor variables are the variables measuring sociodemographic and well-being characteristics.⁵

The canonical discriminant function is a linear combination of the predictor variables (Bibb & Roncek, 1976; Schulman & Cotten, 1993). The canonical correlation coefficient provides information on the discriminating power of the function and is a measure of association that summarizes the degree of relatedness between the groups and the discriminant function. The standardized canonical discriminant function coefficients are measures of the relative importance of each predictor variable; the larger the magnitude of that variable, the greater the contribution of that variable, net the effect of the other predictor variables. The total canonical structure coefficients are analogous to bivariate correlations; they signify how closely a predictor variable and a function are interrelated and are not affected by relationships with other variables.

Discriminant analysis was chosen as the statistical technique to use for several reasons. First, the purpose of this stage of the analysis is to develop an equation that helps to discriminate among the two health information seeking groups; discriminant analysis is a commonly used technique for this purpose (Cleary & Angel, 1984; Huberty & Lowman, 1997). Second, our purpose is not to predict the probability of an event occurring, as would be the case with logistic regression (Morgan et al., 2003); rather we examine a series of variables to see how they combine to significantly discriminate between the two groups of health information seekers. Finally, although logistic regression is often considered superior to discriminant analysis, various researchers have concluded that the results of discriminant analysis and logistic regression are usually very similar (Cleary & Angel, 1984; Fan & Wang, 1999).

⁵Unfortunately, due to smaller *n*'s associated with various skip patterns in the GSS, we were not able to include the Internet usage variables in the discriminant analysis.

Table 1
Characteristics of total GSS sample versus subsample

	GSS 2000 sample	Study subsample
<i>Sociodemographic variables</i>		
Gender		
Female	56%	61%
Ethnicity		
White	84%	83%
Black	13%	12%
Other	3%	5%
Age		
Mean (years)	45	49
Range (years)	18–89	19–89
Total family income		
Modal category	\$25,000 or more	\$25,000 or more
Mean category score	10.71	10.86
Marital status		
Married	58%	49%
Highest degree		
Less than high school	26%	12%
High school	52%	48%
Junior college	4%	8%
Bachelor	13%	22%
Graduate	6%	9%
<i>Internet and computer usage</i>		
Computer use at work		
Yes	61%	62%
Computer use at home		
Yes	97%	97%
Non-Email use of Internet		
Yes	75%	76%
Time spent using Email (hours/week)		
0	40%	34%
1–3	34%	34%
More than 3	26%	33%
<i>Well-being</i>		
Self-reported health status		
Excellent	32%	29%
Good	44%	44%
Fair	19%	19%
Poor	6%	8%
General happiness		
Very happy	32%	31%
Pretty happy	56%	58%
Not too happy	12%	12%
<i>Total respondents</i>	2817	385

Results

Characteristics of the full GSS sample and the subsample of individuals who sought health information and were asked the Internet and health module

questions are included in Table 1. Results show that the subsample used in this analysis is similar in many ways to the full GSS sample of US adults. However, the subsample is slightly more likely to be female and non-married. In addition, they are more likely to have higher levels of education (31% vs. 19% report having a bachelor's degree or higher level of education) and to report spending more time using email than are those in the full sample.

Table 2 compares online vs. offline health information seeking groups. Results indicate that individuals who seek health information online are significantly more likely to be younger (mean age of 40 vs. 52), have higher incomes, and be more educated (48% report having a bachelor's degree or more education compared to 18%) than are offline health information seekers. In addition, online health information seekers are more likely to report using the Internet for purposes other than email, and they spend more time online per week using email. Forty-two percent of online health information seekers report spending more than three hours per week using email, compared to only 15% of offline health information seekers.

Online health information seekers report better well-being when assessed through self-reported health status and general happiness. They are significantly more likely to be healthier and happier than are offline health information seekers. Eighty-six percent of the online group report that their health status was good or excellent, compared to only 60% of the offline group. The offline group is more likely than the online group to report being not too happy.

This comparison suggests that the two groups are very different. These differences are similar to those reflected in the digital divide literature. However, this analysis also reveals that online health information seekers are also healthier and happier than the offline health information group.

Chart 1 shows differences in where members of the two groups seek health information. As the literature suggests, seeking health information from doctors and nurses is highest among both groups. About 80% of each group report using this source. Among the offline group, health magazines, friends or relatives, and television or radio are the next most often cited sources of health information. For the online group, friends and relatives are most often cited, following health care professionals. Fifty-five percent of the online group report seeking health information once or twice via the World Wide Web during the past year. Forty-five percent of this group report using this medium three or more times during the past year.

When comparing sources across the groups, the offline group reports higher usage of all sources except friends and relatives, whereas the online group is slightly more likely to use these sources for health information.

Table 2
Offline versus online health information seeking group characteristics

	Offline health information seekers	Online health information seekers
<i>Sociodemographic variables</i>		
Gender (<i>n</i> = 328)		
Female	61%	64%
Ethnicity (<i>n</i> = 351)		
White	83%	83%
Black	10%	14%
Other	7%	3%
Age (<i>n</i> = 328) ^{***}		
Mean (years)	51.6	40.4
Total family income (<i>n</i> = 328) ^{***}		
Mean category score	9.8	11.2
Marital status (<i>n</i> = 328)		
Married	45%	53%
Highest degree (<i>n</i> = 326) ^{***}		
Less than high school	19%	4%
High school	54%	39%
Junior college	8%	10%
Bachelor	12%	34%
Graduate	6%	14%
<i>Internet and computer usage</i>		
Computer use at work (<i>n</i> = 208)		
Yes	53%	66%
Computer use at home (<i>n</i> = 178)		
Yes	93%	99%
Non-Email use of Internet ^{***} (<i>n</i> = 218)		
Yes	56%	88%
Time spent using email (hours/week) (<i>n</i> = 214) ^{***}		
0	60%	23%
1–3	25%	35%
More than 3	15%	42%
<i>Well-being</i>		
Self-reported health status ^{***} (<i>n</i> = 325)		
Excellent	22%	38%
Good	38%	48%
Fair	26%	9%
Poor	13%	5%
General happiness* (<i>n</i> = 317)		
Very happy	29%	33%
Pretty happy	54%	60%
Not too happy	17%	7%
<i>Health information sources utilized</i>		
Mean number of sources* (<i>n</i> = 325)	3.4	3.8
Total respondents	158	170

p* < 0.05; *p* < 0.01; ****p* < 0.001.

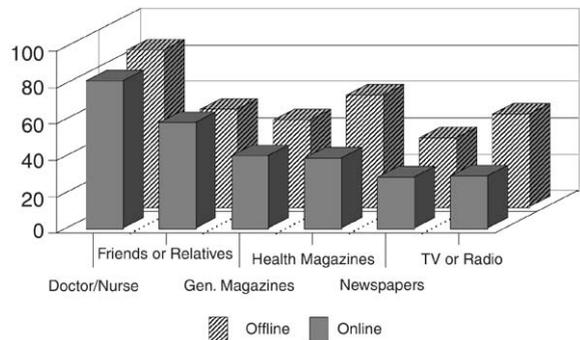


Chart 1. Sources of health information seeking

Offline health information seekers are more likely to seek health information via magazines, both general interest and those related to health. The difference is most dramatic when examining health magazine usage (61% of the offline group, compared to 38% of the online group). The offline group also reports utilizing television and radio significantly more so than the online group (51% compared to 29%).

Table 3 shows the results of the discriminant analysis that uses multiple predictor variables to discriminate among individuals based upon the single dichotomous criterion variable of where they seek health information (online or offline).⁶ This analysis shows that statistically significant differences exist between the two groups. The canonical correlation of 0.523 indicates a fairly strong relationship between the groups and the discriminant function.

Coordinating the signs of the group centroids (means) with those of the standardized discriminant function coefficients allows us to compare the two groups. These coefficients show that age (being older in this case) is associated with the group that sought offline health information. All the other variables in the model distinguish the online health information seeking group from the offline seekers. Age, education, and income make the largest contribution among all the predictor variables (−0.494, 0.463, and 0.423, respectively).

Examination of the total canonical structure coefficients reveals that income, education, age, and health are the dominant variables. Their signs suggest that we might refer to the discriminant function as a general inequality/“digital divide” index. From examining both types of discriminant analysis coefficients, three variables, income, education, and age, stand out as the most important.

Given these results, one can conclude that there are differences between those who engage in online health

⁶A reviewer suggested using logistic regression rather than discriminant analysis. We also ran the model with logistic regression. The results are very similar to those provided in Table 3. Results are available from the authors upon request.

Table 3
Results of canonical discriminant analysis for online and offline health information seekers ($n = 278$)

Predictor variables	Standardized canonical discriminant function coefficients	Total canonical structure coefficients	
Female	0.076	0.092	
White	0.031	0.025	
Age	-0.494	-0.633	
Married	0.014	0.135	
Education	0.463	0.660	
Income	0.423	0.710	
Health	0.222	0.486	
Happiness	0.028	0.288	
Number of sources	0.396	0.318	
Discriminant function statistics			
Canonical correlation	0.523		
Approximate standard error	0.044		
Canonical R^2	0.273		
F Statistic (8 DF)	11.200		
Prob. F	0.0001		
Functions derived	Wilke's lambda	Philla's trace	Prob. F
1	0.727	0.273	0.0001
Group centroids (means)			
	Non-Internet health seekers	-0.643	
	Internet health seekers	0.581	

information seeking vs. those who do not. The main differences between these two groups relate to traditional inequality levels in education, age, and income. Individuals who are older, have lower incomes, and are less educated are less likely to be in the group that uses online health information. These factors help to determine who is likely to seek health information online vs. offline. Previous studies have found that these factors are also related to the “digital divide” (NTIA, 2000).

Conclusions and discussion

As in prior studies (Brodie et al., 1999; Bernhardt, 2001), health care professionals were cited more often than other sources for health information by both the online and offline groups. However, significant segments of both groups relied upon additional sources for health information. Given this finding, a key for researchers is to understand what these sources are and whether social inequalities relate to which resources are being utilized. A key finding of this study is that the results suggest that, indeed, larger societal inequalities discriminate between whether or not individuals utilize online or offline venues for health information.

As we hypothesized, age, income, and education were related to whether individuals sought health information online or offline. These hypotheses were supported at both the bivariate and multivariate levels. At the bivariate level, Internet usage was higher among the Internet health information-seeking group. However, due to data limitations, it was not possible to examine this pattern at the multivariate level. In addition, health was also related, but in a different manner than was hypothesized. At the bivariate level, Internet health information seekers were healthier than were the non-Internet health information seekers. In the discriminant analysis, those who were healthier and happier were less likely to look exclusively offline. This is another key result of this study, which has not been shown in prior research.

Age was a key factor that discriminated between online and offline health information seeking. As these results indicated, the mean age for the non-Internet health information-seeking group was 11-years older than that for the Internet health information seeking group. Given that recent studies show that older adults are one of the groups that are increasingly using the Internet (Fox, 2001), future researchers may find that the strong age differences reported in this study have declined over time. As older adults increasingly use the

Internet for health information seeking, future research will show whether age remains a key discriminating factor in where individuals seek health information.

The unique findings from this study suggest that there are particular factors that discriminate between individuals who seek online vs. offline health information. These factors are ones that have traditionally been associated with the “digital divide” (NTIA, 2000; Bolt & Crawford, 2000). It appears that inequalities associated with education and income result in a lower likelihood of Internet utilization in general, and health information seeking more specifically. Unfortunately, the present data do not allow us to examine whether these patterns exist when controlling for Internet usage. If, as previous research has suggested (see for example, Cotten, 2001; Diaz, 2001; Brodie et al., 2000), using the Internet for health information seeking can help to disseminate health information to marginalized groups and to empower health care consumers more generally, determining ways to increase Internet usage among less educated and lower incomes groups may be one way to decrease inequalities associated with health care provision and decision-making.⁷

The findings regarding health and health information seeking are intriguing. It was hypothesized that individuals who were less healthy would be more likely to use the Internet for health information, as they might be more desperate to find solutions to their health problems. The results of this study suggest that online health information seekers are healthier than the offline group. It may be the case that those who have poorer health status are less physically able to use computers and/or access the Internet; or perhaps health information obtained from the Internet has helped online health seekers to maintain good health. Unfortunately, the data used in the present study do not allow us to examine these relationships.

Given this finding, additional research is needed that examines the interrelationships among health status, “digital divide” factors, and health information seeking. From prior research, we know that individuals who are poorer and less educated are also more likely to have poorer health status (Turner, Lloyd, & Roszell, 1999; Adler & Ostrove, 1999). Longitudinal studies that follow changes in health and digital divide factors over time will help to better ascertain interrelationships and causal patterns among these factors and how they may relate to health information seeking.

Better operationalization of family income, health status, and Internet usage is also needed in future studies. Although the current study controlled for income level, the GSS measure of income was marginal

at best. Response categories were capped at a maximum of \$25,000 or more per year. Given that over 65% of the sample reported income in this category, little is known about the distribution of income levels among these respondents. A further refined measure would yield more information, as would a measure of the number of members in the household. Although the self-report health measure used in the current study has been shown to be reliable and valid in prior studies, more extensive measures assessing a variety of physical and mental health statuses are needed. Given that recent research has shown that particular types of Internet usage are associated with different well-being outcomes (Rohall, Cotten, & Morgan, 2002; Morgan & Cotten, 2003), future studies should also examine a variety of types and levels of Internet usage and how they relate to individuals’ health status and where health information is sought.

Although the results of this study shed new light on factors that distinguish online from offline health information seekers, researchers should note that this study is based on the use of cross-sectional US data. Therefore, issues of causality cannot be determined. Longitudinal research on health information seeking and its relationship to Internet usage is clearly warranted. In addition, due to the manner in which questions are administered via modules in the GSS, only a small subsample of the full GSS sample was administered the set of questions utilized in the current study. Larger representative samples are also needed that examine these issues, as well as studies with international samples that ascertain whether these patterns hold for other populations.

In conclusion, the results of this study provide further insight into differences between individuals who seek health information online and those who employ other means. Aspects related to larger structural inequalities in US society appear to play key roles in discriminating between online and offline health information seeking. A policy implication of these findings is that diminishing these inequalities may be the key to increasing online health information seeking and thus potentially empowering health care consumers.

References

- Adler, N. E., & Ostrove, J. (1999). Socioeconomic status and health: What we know and what we don't. *Annals of the New York Academy of Sciences*, 896, 3–15.
- Beisecker, A. (1988). Aging and the desire for information and input in medical decisions: Patient consumerism in medical encounters. *Gerontologist*, 28, 330.
- Bernhardt, J. M. (2001). Tailoring messages and design in a web-based skin cancer prevention intervention. *International Electronic Journal of Health Education*, 4, 290–297.

⁷Some have suggested that providing assistance in public libraries to teach individuals how to search for information on the Internet would help to close the digital divide.

- Bernhardt, J. M., Lariscy, R. A., Parrott, R. L., Silk, K. J., & Felter, E. M. (2002). Perceived barriers to Internet-based health communication on human genetics. *Journal of Health Communication, 7*, 325–340.
- Bibb, R., & Roncek, D. W. (1976). Investigating group differences: An explication of the sociological potential of discriminant analyses. *Sociological Methods and Research, 4*, 329–379.
- Bischoff, W. R., & Kelley, S. J. (1999). 21st Century house call: The Internet and the World Wide Web. *Holistic Nursing Practice, 13*(4), 42–50.
- Bolt, D., & Crawford, R. (2000). *Digital divide: Computers and our children's future*. New York: TV Books.
- Brodie, M., Flournoy R. E., Altman, D. E., Blendon, R. J., Benson, J. M., & Rosenbaum, M. D. (2000). Health information, the Internet, and the Digital Divide. *Health Affairs, 19*, 255–265.
- Brodie, M., Kjellson, N., Hoff, T., & Parker, M. (1999). Perceptions of Latinos, African Americans, and Whites on media as a health information source. *The Howard Journal of Communication, 10*, 147–167.
- Brodie, M., Foehr, U., Rideout, V., Baer, N., Miller, C., Flournoy, R., & Altman, D. (2001). Communicating health information through the entertainment media. *Health Affairs, 20*, 192–199.
- Cleary, P. D., & Angel, R. (1984). The analysis of relationships involving dichotomous dependent variables. *Journal of Health and Social Behavior, 25*(3), 334–348.
- Cline, R., & Haynes, K. (2001). Consumer health information seeking on the Internet: The state of the art. *Health Education Research, 16*(6), 671–692.
- Cotten, S. (2001). Implications of Internet technology for medical sociology in the new millennium. *Sociological Spectrum, 21*, 319–340.
- Davis, J. J. (2002). Disenfranchising the disabled: The inaccessibility of Internet-based health information. *Journal of Health Communication, 7*, 355–367.
- Diaz, M. B. (2001). From the waiting room to the computer terminal: Internet access changes patient–doctor relationships. *Hispanic, 14*(3), 56.
- Elliott, B., & Elliott, G. (2000). High volume medical web sites. *Delaware Medical Journal, 72*(1), 21–29.
- Eng, T., Maxfield, A., Patrick, K., Deering, M. J., Ratzan, S., & Gustafson, D. (1998). Access to health information and support: A public highway or a private road? *Journal of the American Medical Association, 280*(15), 1371–1375.
- Eysenbach, G., & Kohler, C. (2002). How do consumers search for and appraise health information on the World Wide Web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal, 324*(7337), 573–577.
- Fan, X., & Wang, L. (1999). Comparing linear discriminant function with logistic regression for the two-group classification problem. *Journal of Experimental Education, 67*(3), 265, 22p.
- Fattah, H. (2000). Click here for health. *American Demographics, 22*(12), 12–14.
- Fox, S. (2001). Wired Seniors. Pew Internet & American Life Project: Online Report. Retrieved September 20, 2001 from <http://www.pewinternet.org/reports/pdfs/PIP.Wired.Seniors.Report.pdf>.
- Fox, S., & Fallows, D. (2003). Internet health resources: Health searches and email have become more commonplace, but there is room for improvement in searches and overall Internet access. Pew Internet & American Life Project: Online Report. Retrieved July 18, 2003 from <http://www.pewinternet.org/reports/toc.asp?Report=95>.
- Fox, S., & Rainie, L. (2000). The online health care revolution: How the Web helps Americans take better care of themselves. Pew Internet & American Life Project: Online Report. Retrieved November 28, 2000 from <http://www.pewinternet.org/reports/pdfs/PIP.Health.Report.pdf>.
- Fox, S., & Rainie, L. (2002). Vital decisions: How Internet users decide what information to trust when they or their loved ones are sick. Pew Internet & American Life Project: Online Report. Retrieved June 10, 2002 from <http://www.pewinternet.org/reports/pdfs/PIP.Vital.Decisions.May2002.pdf>.
- Gallagher, S. M. (1999). Rethinking access in an information age. *Ostomy Wound Management, 45*(9), 12–14, 16.
- Giddens, A. (1991). *Modernity and self-identity*. Cambridge: Polity Press.
- Goldsmith, J. (2000). How will the Internet change our health system? *Health Affairs, 19*(1), 148–156.
- Grandinetti, D. A. (2000). Doctors and the Web: Help your patients surf the Net safely. *Medical Economics, 63*(8) April, 28–34.
- Hardey, M. (1999). Doctor in the house: The Internet as a source of lay health knowledge and the challenge to expertise. *Sociology of Health and Illness, 21*(6), 820–835.
- Harris Interactive. (2001). eHealth traffic critically dependent on search engines and portals. Retrieved November 4, 2001 from <http://www.harrisinteractive.com/news/allnewsbydate.asp?NewsID=270>.
- Hern, M. J., Weitkamp, T., Hillard, P. J. A., Trigg, J., & Guard, R. (1998). Promoting women's health via the World Wide Web. *Journal of Obstetric, Gynecologic, & Neonatal Nursing, 27*(6), 606–610.
- Huberty, C. J., & Lowman, L. L. (1997). Discriminant analysis via statistical packages. *Educational and Psychological Measurement, 57*(5), 759–784.
- Impicciatore, P., & Pandolfini, C. (1997). Reliability of health information for the public on the World Wide Web: Systematic survey of advice on managing fever in children at home. *British Medical Journal, 314*(7097), 1875–1879.
- Jadad, A. R., & Gagliardi, A. (1998). Rating health information on the Internet: Navigating to knowledge or babel? *Journal of the American Medical Association, 279*(8), 611–614.
- Kakai, H., Maskarinec, G., Shumay, D. M., Tatsumura, Y., & Tasaki, K. (2003). Ethnic differences in choices of health information by cancer patients using complementary and alternative medicine: An exploratory study with correspondence analysis. *Social Science & Medicine, 56*, 851–862.
- Kommers, N., & Rainie, L. (2002). Use of the Internet at major life moments. Pew Internet & American Life Project: Online Report. Retrieved May 15, 2002 from <http://www.pewinternet.org/reports/pdfs/PIP.Major.Moments.Report.pdf>.
- Kummervold, P. E., Gammon, D., Bergvik, S., Johnsen, J. K., Hasvold, T., & Rosenvinge, J. H. (2002). Social support in a wired world: Use of online mental health forums in Norway. *Nordic Journal of Psychiatry, 56*(1), 59–65.
- Kurtz, M. E., Kurtz, J. C., Johnson, S. M., & Cooper, W. (2001). Sources of information on the health effects of

- environmental tobacco smoke among African-American children and adolescents. *Journal of Adolescent Health*, 28(6), 458–464.
- McLeod, S. D. (1998). The quality of medical information on the Internet: A new public health concern. *Archives of Ophthalmology*, 116(12). Retrieved July 12, 2000 from <http://archophth.amaassn.org/issues/v116n12/full/eed8293.html>.
- Montagne, M. (2001). Mass media representations as drug information for patients: The Prozac phenomenon. *Substance Use & Misuse*, 36, 9–10.
- Morgan, C., & Cotten, S. R. (2003). The relationship between Internet activities and depressive symptoms in a sample of college freshmen. *CyberPsychology & Behavior*, 6(2), 133–142.
- Morgan, G. A., Vaske, J. J., Gliner, J. A., & Harmon, R. J. (2003). Logistic regression and discriminant analysis: Use and interpretation. *Journal of the American Academy of Child Adolescent Psychiatry*, 42(8), 994–997.
- National Telecommunications and Information Administration (NTIA) Annual Report (2000). Retrieved February 10, 2001 from <http://www.ntia.doc.gov/ntiahome/annualrpt/2001/2000annrpt.htm>.
- Nielsen//NetRatings. (2002). Health and finance sites capture fastest growing rankings in January. Retrieved March 1, 2003 from http://www.nielsen-netratings.com/pr/pr_020203.pdf.
- Pandey, S., Hart, J., & Tiwary, S. (2002). Women's health and the Internet: Understanding emerging trends and implications. *Social Science & Medicine*, 56, 179–191.
- Patrick, K. (2000). Information technology and the future of preventive medicine: Potential, pitfalls, and policy. *American Journal of Preventive Medicine*, 19, 132–135.
- Pennbridge, J., Moya, R., & Rodrigues, L. (1999). Questionnaire survey of California consumers' use and rating of sources of health care information including the Internet. *Western Journal of Medicine*, 17(5-6), 302–305.
- Pirisi, A. (2000). Low health literacy prevents equal access to care. *The Lancet*, 356, 1828.
- Preece, J. (2000). *Online communities: Designing usability, supporting sociability*. New York: Wiley.
- Rainie, L., & Packel, D. (2001). More online, doing more. Pew Internet & American Life: Online Report. Retrieved February 27, 2001 from http://www.pewinternet.org/reports/pdfs/PIP_Changing_Population.pdf.
- Rohall, D., Cotten, S. R., & Morgan, C. (2002). Internet use and the self concept: Linking specific uses to global self-esteem. *Current Research in Social Psychology*, 8(1), 1–18.
- Schulman, M., & Cotten, S. (1993). Adaptations to the farm crisis: Macro level implications of micro level behaviors. *Applied Behavioral Science Review*, 1, 93–111.
- Shackley, P., & Ryan, M. (1994). What is the role of the consumer in health care? *Journal of Social Policy*, 23(4), 517–536.
- Skinner, H., Biscope, S., & Poland, B. (2003). Quality of Internet access: Barrier behind Internet use. *Social Science & Medicine*, 57, 875–880.
- Spooner, T., & Rainie, L. (2000). African-Americans and the Internet. Pew Internet & American Life: Online Report. Retrieved June 10, 2002 from http://www.pewinternet.org/reports/pdfs/PIP_African_Americans_Report.pdf.
- Suarez-Amazon, M. E., Kendall, C. J., & Dorgan, M. (2001). Surfing the Net—Information on the World Wide Web for persons with arthritis: Patient empowerment or patient deceit? *Journal of Rheumatology*, 28(1), 185–191.
- Tardy, R. W., & Hale, C. (1998). Getting “plugged in”: A network analysis of health-information seeking among ‘stay-at-home moms’. *Communication Monographs*, 65(4), 336–357.
- Taylor, M. R. G., Alman, A., & Manchester, D. K. (2001). Use of the Internet by patients and their families to obtain genetics-related information. *Mayo Clinic Proceedings*, 76(8), 772–776.
- Turner, R. J., Lloyd, D. A., & Roszell, P. (1999). Personal resources and the social distribution of depression. *American Journal of Community Psychology*, 27(5), 643–672.
- Zarcadoolas, C., Blanco, M., Boyer, J. F., & Pleasant, A. (2002). Unweaving the Web: An exploratory study of low-literate adults' navigation skills on the World Wide Web. *Journal of Health Communication*, 7, 309–325.