Effects of internet use on health and depression: a longitudinal study.

Katie Bessière  
*Carnegie Mellon University*

Sarah Pressman  
*University of Kansas*

Sara Kiesler  
*Carnegie Mellon University*

Robert Kraut  
*Carnegie Mellon University*

Follow this and additional works at: [http://repository.cmu.edu/hcii](http://repository.cmu.edu/hcii)

Part of the [Computer Sciences Commons](http://repository.cmu.edu/hcii)

Published In  
*Journal of medical Internet research, 12, 1, 6-6.*
Effects of Internet Use on Health and Depression: A Longitudinal Study

Reviewed by Sarah Bass
Katie Bessiere, MS,† Sarah Pressman, PhD,‡ Sara Kiesler, PhD,† and Robert Kraut, PhD§

Abstract

Background

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

Objective

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

Methods

Data come from a national US panel survey of 750 individuals conducted from 2000 to 2002. Across three surveys, respondents described their use of the Internet for different purposes, indicated whether they had any of 13 serious illnesses (or were taking care of someone with a serious illness), and reported their depression. In the initial and final surveys they also reported on their physical health. Lagged dependent variable regression analysis was used to predict changes in depression and general health reported on a later survey from frequency of different types of Internet use at an earlier period, holding constant prior depression and general health, respectively. Statistical interactions tested whether uses of the Internet predicted depression and general health differently for people who initially differed on their general health, chronic illness, and caregiver status.

Results

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

Conclusions

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

Keywords: Depression, health, social support, Internet, longitudinal survey

Introduction

The rapid expansion of the Internet has greatly increased the amount of health information available to the general public. In 2001, approximately 43% of the US public was using the Internet, and an estimated 40% reported using Internet health resources in the previous year [1]. More recently, in 2009, about 74% [2] of US adults said they used the Internet. While email and general information searches remain the most common uses of the Internet [3], using the Internet to obtain health resources is also very popular [4], with over 50% of Internet users searching for medical information [4]. With millions using the Internet, and a large proportion of the population explicitly using it for health purposes, it is important to assess how this particular use of the Internet is affecting people’s well-being, especially their physical and mental health.

Although prior research has shown that use of the Internet to communicate with friends and family is associated with declines in depression [4], little reliable information exists about the impact of using the Internet to obtain health resources, especially in nonclinical populations [5-8]. The current study used data from a national US random household sample survey to address the impact of Internet use to obtain health information and support on well-being and health. We also examined whether these effects differed for people with differing levels of health and caregiver status, and whether these well-being associations were present for other types of Internet use.

Health Resources Online

Traditionally, physicians held and filtered health information for the general public [6]. Given the easy availability of health information online, people can now bypass medical professionals entirely and find information, advice, support, and even treatments (eg, pharmaceuticals and herbal remedies) on the Internet [6]. The Internet contains thousands of medically relevant websites such as large information sites (eg, WebMD and Medline), disease support groups, discussion boards, distributions lists, personal websites, and websites selling “miracle cures”. Online communities and support groups allow people to talk about their health problems with others and are popular both as a source of social support and as a primary source of information [10-12]. Online communities and support groups also provide support to caregivers and the families of people with serious health problems [13].

Much of the research examining the effects of these online health resources has examined its use by groups with a specific illness or disease [11]. Although some writers have warned about poor quality medical information online [8] and patients’ inability to distinguish poor advice from good advice [15], there are few reports of serious harm [16]. Both qualitative studies [18] and polls assessing consumers’ beliefs [15,17] suggest that online health information and health support can improve patients’ and caregivers’ optimism and feelings of control. For example, 32% of respondents who used one consumer health site reported their condition had improved after visiting the site, 64% said they felt they were better informed, and 50% said they believed that the information they learned changed how they felt about their condition [19]. These types of retrospective data are unreliable on many dimensions, including that typically do not compare Internet users’ reports of health benefits to those of non-users. However, they are one source for the popular belief that use of the Internet for health purposes can improve well-being. These self-report studies are supported by clinical trials of online interventions, which suggest that these programs can reduce patients’ pain and stress and improve their physical performance and perceived social support [9,19,20].

Many studies of online health resources include the caregivers of the chronically ill in the sample [21] because of the high levels of stress experienced by this population as well as their strong interest in the health of those they care for [22]. Overall, little is known about the extent to which individuals with health problems (or caregivers of the ill) benefit from having unfettered access to online health information and support. The literature cited above would lead one to believe that online health...
When choosing the optimal treatment strategy for Internet users, it is important to consider the potential benefits and risks associated with each approach. For instance, while Internet use may provide a sense of connection and social support, it can also lead to increased social isolation and depression. Similarly, while Internet use may serve as a source of information and entertainment, it can also result in increased anxiety and stress.

To address these concerns, we conducted a study to assess the impact of Internet use on well-being. Our sample consisted of 2700 adults who agreed to participate in the study. We also oversampled on Internet users and compared them to the 2000 US Census data. Our results indicate that Internet users are younger and wealthier than the general population, and more likely to be Caucasian and female.

The purpose of this study was to determine whether using the Internet for health purposes is beneficial or harmful to physical and psychological well-being. We were also interested in whether this association would be moderated by people's health or care-giving status. Specifically, we examined whether individuals with a good reason for searching the Internet for medical information would fare differently from their healthy (or non care-giving) counterparts. We were also interested in whether this association would be moderated by people's health or care-giving status. Specifically, we examined whether individuals with a good reason for searching the Internet for medical information would fare differently from their healthy (or non care-giving) counterparts.

We conducted a follow-up survey to assess the impact of Internet use on well-being. The follow-up survey was conducted 12 to 18 months after the initial survey. Our results indicate that Internet users are more likely to be younger, wealthier, and more likely to be Caucasian and female. The majority of Internet users were male (53%), and 89% were Caucasian. The median household income of Internet users was US $30-50,000, which is significantly higher than the median household income of the general population (US $41,900).

The implications of our findings are significant. Internet users are more likely to be younger, wealthier, and more likely to be Caucasian and female. This suggests that Internet use may be particularly beneficial for people who are engaged in social and leisure activities, and may provide a sense of connection and social support. However, Internet use may also lead to increased social isolation and depression. Therefore, it is important to consider the potential benefits and risks associated with each approach, and to develop strategies to address these concerns.
measures and age = -.26). The outcome measures were relatively stable (correlation between general health at times 1 and 3, \( r = -.35 \)).

### Preliminary Analyses

Preliminary Analyses suggested that those with higher levels of depression face worse health outcomes even if they are not clinically depressed \([58,59]\). Second, a continuous measure of depression is more sensitive than a dichotomous one, thereby reducing the type II error of failing to identify relationships that actually exist. Respondents described how frequently they experienced various symptoms of depression in the past week, with \( r = 0 \) days with a symptom, and \( r = 4 \) experienced the symptom 5 to 7 days in the preceding week (Cronbach alpha = .89).

### General Health

The CES-D was designed to assess the degree of depression in the general population. We used the CES-D as a continuous measure of degree of depression rather than using a cut-off to classify participants as clinically depressed for two reasons. First, prior research indicates that those with higher levels of depression face worse health outcomes even if they are not clinically depressed \([58,59]\). Second, a continuous measure of depression is more sensitive than a dichotomous one, thereby reducing the type II error of failing to identify relationships that actually exist. Respondents described how frequently they experienced various symptoms of depression in the past week, with \( r = 0 \) days with a symptom, and \( r = 4 \) experienced the symptom 5 to 7 days in the preceding week (Cronbach alpha = .89).

### General Health

We asked the one-item general health question from the 36-item Medical Outcome Study Short-Form Health Survey (MOS SF-36) \([60,61]\). The item asked, "In general, would you say your health is..." scored from 1 = poor to 5 = excellent. Because of an error, this item was omitted from the second survey and measured only on the first and third ones.

### Statistical Analysis

To test hypotheses about changes in participants' general health and levels of depression, we used regression analysis with a lagged dependent variable, as recommended by Cohen et al \([62]\). This analysis predicts participants' levels of general health or depression at a later time period from variables measured at an earlier time period, including control variables, measures of Internet use, and the lagged dependent variable (ie, their prior health or depression score). Since the health outcome was assessed only on the first questionnaire (at time 1) and on the second questionnaire (at time 3), there is only a single lagged dependent variable and thus a single observation per respondent; changes in general health, therefore, were analyzed using ordinary least-squares regression. However, depression was assessed on all three questionnaires (at times 1, 2, and 3). Therefore, there are two lagged dependent variables measuring depression (depression at time 2 controlling for depression at time 1 and depression at time 3 controlling for depression at time 2). This resulted in two observations per respondent. To adjust both the coefficients and standard errors for the nonindependence of observations within respondents, we analyzed changes in depression using random effects regression, with the respondent as the random effect. Because prior levels of the outcome variables (ie, health and depression) are included in the analyses, both the dependent variables and the other predictor variables have been adjusted for the initial levels of health or depression. Lagged dependent variable regression is appropriate for testing dynamic theories of change, in which states or events at one time influence states or events at a later time \([63]\). Lagged dependent variable models are appropriate when the data exhibit "stationarity," that is, the dependent variable at each time period comes from the same population with a common mean and variance, and the serial correlation in the residuals is not high (ie, less than .50). Preliminary tests indicated that these conditions were met, making lagged dependent variable regression appropriate for the research reported here.

The sample includes all respondents who had change scores on the depression and health outcomes: 916 respondents who completed the CES-D depression index at least twice and 671 who described their general health twice. We included in the analyses participants who had never used the Internet. Our analyses did not change when we eliminated respondents who had never used the Internet. Some respondents omitted one or more independent variables (such as their initial health status, income, or estimates of Internet use), causing a reduction in sample size in the regressions and potential biases in results if the respondents with missing data were not a random subset of the full sample. To correct for these problems, we used multiple imputation to fill in missing values, as suggested by Rubin \([64]\). Multiple imputation creates multiple datasets that estimate missing independent variables from the nonmissing data. These data are then analyzed using conventional statistical methods and combined to produce estimates and confidence intervals that incorporate the uncertainty resulting from imputing missing data \([65]\).

### Results

#### Preliminary Analyses

Descriptive statistics for the sample and correlations among variables are presented in Multimedia Appendix 1. The correlations showed moderate associations among all the Internet use variables (mean \( r = .50 \)). In addition, older respondents were less likely to use the Internet than younger ones (mean \( r \) between various Internet use measures and age \( r = -.26 \)). The outcome measures were relatively stable (correlation between general health at times 1 and 3, \( r = .69 \) and between depression at times 1 and 3, \( r = .59 \)). The measures of general health and depression were themselves moderately negatively correlated (mean \( r \) at a single time period \( r = -.53 \)).
We conducted a preliminary regression analysis using demographic controls and measures of initial health status to predict the six different components of Internet use at time 1. When predicting use of the Internet for a specific purpose we included the other components of Internet use as predictor variables to control for overall propensity to use the Internet. These analyses are summarized in Tables 1 and 2. As shown in Models 1A to 1F, at the time of the first survey, women were more likely to use the Internet to communicate with friends and family online and for health purposes, whereas men were more likely to use the Internet for information and escape. Younger participants reported more use of the Internet to meet people and participate in online groups and for information and entertainment, whereas older participants reported more use of the Internet for health purposes. Married respondents were less likely than single respondents to use the Internet to communicate with friends and family and to meet new people. More highly educated respondents were more likely to use the Internet to communicate with friends and family, for information, and for shopping, but less likely to use it for entertainment. Wealthier respondents were more likely to use the Internet to communicate with friends and family, for information, and for shopping, but less likely to use it for health purposes. These demographic results echo results of national polls. 

An alternative interpretation is that people who choose to seek out Internet health resources may be especially sensitive to hypochondriasis or excessive worry about minor physical health, and whether these consequences were comparable to consequences of other uses of the Internet. We examined whether the impact of using the Internet to obtain health resources was associated with increased depression after accounting for initial depression. Using the Internet to obtain health resources was associated with increased depression after accounting for initial depression ($P = .002$). People who used the Internet for communication with friends and family, on the other hand, reported reduced depression after accounting for initial depression ($P = .002$). The changes in depression associated with Internet use were small. Increasing use of the Internet for health purposes by one unit (eg, from 3 to 5 days per week to once a day) was associated with 11% of a standard deviation increase in depression. Conversely, a unit increase in Internet use to communicate with friends and family was associated with 7% of a standard deviation decrease in depression. Using the Internet for meeting new people, escape, entertainment, or shopping was not associated with changes in depression.

Of particular interest to the present research is that at the time of the first questionnaire, respondents who had a serious illness themselves or those who cared for someone with a serious illness were more likely to use the Internet for health purposes. Those who were more depressed at the time of the first questionnaire were more likely to use the Internet for escape and to obtain health resources, but were less likely to use it for communicating with friends and family or for shopping.

### Changes in Depression and Health

Table 3 summarizes the lagged dependent variable regression analysis of CES-D depression at a later time from demographic variables and the types of Internet use at an earlier time, controlling for depression at the earlier time. The intercept in this model is the expected level of depression on the second and third questionnaires among nonwhite, unmarried females who did not have a serious illness themselves nor were they caring for someone who did, and who had average levels of education, income, health, depression, and Internet use at the first questionnaire. Those who were more depressed at the initial period were also more depressed at the subsequent period ($P < .001$). Being younger, wealthier, and male were all associated with reductions in symptoms of depression after accounting for initial depression. Using the Internet to obtain health resources was associated with increased depression after accounting for initial depression ($P = .002$). People who used the Internet for communication with friends and family, on the other hand, reported reduced depression after accounting for initial depression ($P = .002$). The changes in depression associated with Internet use were small. Increasing use of the Internet for health purposes by one unit (eg, from 3 to 5 days per week to once a day) was associated with 11% of a standard deviation increase in depression. Conversely, a unit increase in Internet use to communicate with friends and family was associated with 7% of a standard deviation decrease in depression. Using the Internet for meeting new people, escape, entertainment, or shopping was not associated with changes in depression.

### Predicting depression from Internet use

We anticipated that respondents’ initial health and caregiver status would moderate the association between Internet use and subsequent depression and health. We tested this expectation by adding to the models presented in Tables 3 and 4 interactions of the six types of Internet use with the dummy variables representing whether the respondent had a serious illness or took care of a household member with a serious illness at the earlier time period. None of the interactions of Internet use with prior illness or caregiver status was significant (all $Ps > .20$; see Multimedia Appendix 2 and Multimedia Appendix 3 for complete results). Thus, these interaction analyses provided no evidence that the impact of using the Internet for health purposes influenced well-being in different ways for those who had a serious illness or who cared for someone with a serious illness compared to a healthy or noncaretaking sample.

### Discussion

The goals of this study were to determine whether people’s use of the Internet to obtain health resources would have consequences for their psychological well-being and physical health, and whether these consequences were comparable to consequences of other uses of the Internet. We examined whether the impact of using the Internet to obtain health resources might be moderated by participants’ initial health or caregiver status. We found that using the Internet to obtain health information was associated with increased depression over approximately 6 to 8 months, while using it to communicate with friends and family was associated with decreased depression. Interestingly, these associations did not depend on the initial health status of the participants (eg, the presence of serious illness) or whether they were the primary caregiver for an ill person. Furthermore, these uses of the Internet were not associated with changes in individuals’ ratings of their general health.

We did not expect that using the Internet for health purposes would be associated with increases in depression. There are, however, several plausible explanations for this finding. First, the Internet has both good and poor quality medical advice that is difficult to for an untrained observer to distinguish. For example, a previous study revealed that only 20% of websites provided correct information on how to take a child’s temperature. Furthermore, only one third of users verify Internet information with their doctor to ensure accuracy. It may be that one source of the increase in depression is the misinformation people get from factually incorrect websites. This may lead to inaccurate self-diagnosis, poor health behaviors (eg, herbal remedies), or potentially unnecessary worry (for both healthy and ill populations). This negative rumination could occur when researching one’s own medical problems or those of loved ones, easily leading to depressive symptoms at a later point in time.

Another possible source of depression may come about when people use online health support groups. While health support groups moderated by doctors, nurses, or trained moderators may be an important source of health information, many online support groups do not have professional moderation, and most are composed of strangers. Both the information and the empathy and other types of emotional support people receive from strangers they meet in online support groups may be less valuable than the resources they could get from offline interactions with family and friends. The advice offered in support groups often consists of personal anecdotes that may not be as helpful as medically relevant information. Moreover, too much time spent in online support groups may displace in-person support and as a result harm the psychological well-being people derive from interaction with friends and family offline. Results from the current study showing that communicating online was associated with declines in depression when the communication was with friends and family but not with new people; and results from prior research showing that communication with strangers online may lead to increased depression suggest that discussion of health problems with strangers online may be problematic.

An alternative interpretation is that people who choose to seek out Internet health resources may be especially sensitive to hypochondriasis or excessive worry about minor
health symptoms or perceived health risks. The association between baseline depression and seeking health information online is evidence of this. Internet websites and support groups might be compelling for such persons, as they are rife with lists of symptoms, narratives of pain and grief, dire warnings about treatment side effects, and even graphic photos of diseased organs. Reading about symptoms and anecdotes from patients may cause this group to imagine being ill and to inflate their perceptions of risk. Information and discussions of health problems also may cause them to ruminate [28-30] and increase their anxiety [27-29]. Consistent with this argument is evidence suggesting that those with psychosomatic illnesses are particularly likely to use Internet health resources [23,75], and that people with high levels of health anxiety or hypochondriasis use health resources significantly more than their nonanxious counterparts.

The finding that online communication with friends and family reduced the frequency of depression symptoms is consistent with a large literature on social support [76,77] and warrants little further discussion. If online communication with friends and family increases perceived social support, this could lead to lower depression and improved psychological well-being. Similarly, maintaining contact with friends and family may enhance the quality of relationships, decrease loneliness and social isolation, and improve the nature of the social network, all of which have been tied to lower depression and improved psychological well-being [28-81].

Although one might be concerned about regression toward the mean, this statistical artifact cannot account for results showing that use of the Internet for health information was associated with increases in depression, but use for communication with friends and family was associated with decreases. Indeed regression artifacts would have produced a pattern of results opposite to the ones reported here. Regression toward the mean occurs because of measurement error, when error causes extreme scores at one measurement period to be less extreme at a different period (eg, people who reported many symptoms of depression at the first survey should report fewer at subsequent periods). Models 1A and 1F in Tables 1 and 2 show that people who were initially more depressed were more likely to use the Internet for health purposes and less likely to use it to communicate with friends and family. If their initial depression caused this pattern of Internet use, then regression toward the mean would cause those people who were initially severely depressed (and therefore selected to use the Internet for health purposes) to appear less depressed at the later period.

The absence of significant effects of the Internet on changes in participants' general health may stem from insensitivity of the health measure. First, we used a single-item measure of general health, which limits our assessment to perceptions of general health as opposed to any one specific symptom. It is also possible that there is no link between physical health and Internet health resource use, or that it is weakened by other behaviors. For example, if using the Internet for health purposes resulted in appropriate treatment seeking, this might mask a possible connection. Unfortunately, since we do not have data on actions taken or health center visits, we cannot determine if this was the case. Thus far, however, the literature does not show such effects, at least not for unevaluated health resources and unsupervised use [82]. Most previous studies of online health resources have measured attitudes rather than behavior change. The few studies examining behavior change [83] have been clinical trials involving physician-supervised, closed Internet sites or moderated support groups rather than free access to Internet resources.

Limitations

The longitudinal data and analyses reported here allow stronger causal claims about the relationship between Internet use and depression than do cross-sectional data. Because the same individuals were measured at multiple time points, stable characteristics such as demographic differences and personality are automatically controlled when assessing changes in depression. In addition, the lagged dependent variable analysis controls for initial levels of depression and health when predicting subsequent levels of depression and health. Despite the benefits of longitudinal analyses, however, we cannot establish causality solely based on them. In particular, longitudinal analyses do not control for unmeasured variables, such as hypochondriasis, stress, or health behaviors that may also change over time and predict both changes in depression and changes in use of the Internet for health resources. Clinical trials with random assignment are needed to make stronger causal claims.

Our ability to generalize to the US population is limited because we over-sampled Internet users and because only 35% of those initially contacted by telephone completed an outcome measure at least two times. Finally, our data were collected during 2000 to 2002, and the Internet and access to it have changed dramatically since then. The quality of health information and support online may have improved, and Internet users today may no longer use Internet resources in the same fashion as they did during the time period of our study, suggesting the need for follow-up research in this area.

Conclusion

This study examined the consequences of using Internet health-related resources in a way that other studies have not. We used a national sample including both healthy and ill people, and administered a longitudinal survey to discover if using the Internet to obtain health information or interpersonal communication with friends and family predicted subsequent changes in participants' reported depression and general health. Our results suggest that using the Internet to obtain health resources is associated with increases in symptoms of depression. This finding cannot be interpreted as a broad effect of being online, since we also showed that communicating online with friends and family was associated with declines in symptoms of depression. However, since we did not control the uses of the Internet chosen by respondents, we cannot determine whether these effects were due to characteristics of the individuals or the nature of the online resources they used, or both. Additional research is needed to determine what leads individuals to seek out health resources online, whether the information that they discover (and believe) is factually correct, and what actions ensue.

Acknowledgments

This research was supported by NSF grant IIS-9900449 and a grant from the Pew Charitable Trust.

Multimedia Appendix 1

Means/percentages and correlations among variables

Click here to view. (8.0K, pdf)

Multimedia Appendix 2

Intervention Characteristics for Interventions Included in the Meta-Analysis

Click here to view. (208K, pdf)

Multimedia Appendix 3

Predicting health from respondents' prior use of the Internet and interactions with health and caregiver status

Click here to view. (208K, pdf)

References


