**Internet’s Effect on Healthcare Consumption**

by

Sachi Shetty

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Business and Economics Honors Program

NYU Shanghai

May 2022

Professor Marti G. Subrahmanyam Dean Sherry Glied

Professor Christina Wang

Professor Wendy Jin

Faculty Advisers Thesis Adviser

**Table of Contents**

**Abstract 2**

**Preface 3**

**Chapter 1. Introduction 4**

1.1 Literature Review 6

1.2 Methods and Methodology 10

**Chapter 2. Results 15**

**Chapter 3. Discussion 20**

**References 22**

**Acknowledgments 24**

**Abstract**

Since the turn of the millennium, the use of online health information has only become more prevalent. This technology has the potential to make medical information more accessible and give individuals more autonomy in the way they consume healthcare. It also has the potential to reduce inefficiencies in the use of healthcare appointments. This capstone project aims to answer the question: how has readily available online health information altered the way individuals consume healthcare? Existing literature suggests that individuals use online health information to inform, drive, and supplement the existing advice of their healthcare providers. There has also been much concern over the growth of ‘cyberchondria’, excessive worrying about one's health that has been fueled by the influx of health information. This study used two datasets, to construct a propensity to use the internet for health information scores and then compare it to the efficiency of the given ‘reason for visit’. The study found that individuals who use the internet as a source of health information are less inclined to use medical appointments effectively. This capstone project welcomes further research into understanding

**Keywords:** Internet Health Information, Patient–health professional relationship, Resources, Health literacy, Effectiveness of Care

**Preface**

The motivation for researching the impact of online health information stems from my own experience. As a young adult, I am constantly turning to the internet for health information when debating whether or not a symptom warranted medical attention. I have also noticed a shift in the way physicians have begun to refer patients to online resources and was interested in understanding the dynamics of the shift in healthcare consumption.

With the help of my advisor Dean Sherry Glied, I have conducted quantitative research using two prominent US health surveys. I was able to draw conclusions on the impact of online health searches on the use of efficient medical care, which has real-world implications.

**Chapter 1. Introduction**

The considerable asymmetry of information between doctors and patients is shrinking due to the growing prevalence of online symptom checkers such as MayoClinic and WebMD.[1](https://www.zotero.org/google-docs/?2qfo9K) Health information is now easily accessible to those with an internet connection. Internet sources can create a more informed use of healthcare and improve the health of those who cannot afford unnecessary doctor visits. In countries like the United States, where healthcare costs are climbing, the internet gives people a tool to assess whether or not their injuries or illness can benefit from visiting a health professional.[2](https://www.zotero.org/google-docs/?Imb66J) This development has the potential to create more efficient use of the healthcare system. This study however concluded that individuals who used the internet as a source of health information were less inclined to use medical appointments effectively.

The newly gained access to health information can empower individuals and shift how individuals interact with the healthcare system. By having access to information on their health condition(s) online, patients can be conscious of their healthcare options. They can become more central in the decision-making process surrounding their health.[3](https://www.zotero.org/google-docs/?jAeDzx) For example, in the case of type 2 diabetes, diet and exercise (lifestyle changes) can influence health.[4](https://www.zotero.org/google-docs/?YaN4uq) The ability to search what foods are suitable for you instantly might be more impactful than a quick discussion with your doctor months before on avoiding refined carbohydrates.

However, having this much access to health information also comes with its downsides. Two main risks have emerged with individuals becoming more autonomous in their health decisions. The first is a phenomenon called cyberchondria—a term coined by combining ‘cyber’ and ‘hypochondria’, a health anxiety disorder—which has gained ground since the dissemination of the internet.[5](https://www.zotero.org/google-docs/?En9TfH) Second, individuals may be inaccurately self-diagnosing underplaying the nature of their illness and avoid seeking healthcare in cases where there could be advantages to medical intervention. This distortion of information can lead to an over/underconsumption of healthcare that was not as prevalent before the internet.

The motivation of this project is to investigate how this readily available online health information has altered the way individuals consume healthcare, and how healthcare consumption will continue to adapt over time. Who are the people that use the internet to decide whether or not to visit a health professional, and are their decisions more informed? Does the internet allow individuals to use the healthcare system more efficiently or is it counterproductive? Are people able to determine how serious their illness is accurately based on the internet, or is access to too much health information causing people to over-consume healthcare inefficiently?

I hypothesize that individuals who are already health-conscious will be the ones to utilize online health information to research their concerns. I assume those who trust the internet as a source of information are more likely to use online sources as a litmus test to decide whether or not to visit a healthcare provider. I presume this trend will also be higher among younger individuals, and that it will have become more common practice overtime to bring health concerns and prior research to doctor appointments. I believe I will find that individuals are coming to the doctor's office with more targeted health concerns and using appointments more effectively (ie. leave with a prescription, lab order, or other forms of direct follow-up treatment plan). I hypothesize there will be a positive correlation between internet usage and the effectiveness of a medical appointment.

To investigate the impact of online health information on the effectiveness of a medical visit, two datasets were used. The first dataset, the Health Information National Trends Survey, was used to obtain the demographic and socioeconomic variables correlated to the use of online health information and generate a profile on who uses the internet for medical searches. This individual’s propensity to use the internet for health information was then applied to the second dataset, the National Ambulatory Medical Care Survey, which detailed the actual nature of each medical appointment. The “Reasons for Visit” were ranked based on efficiency and this efficiency score was compared to an individual's propensity to use online health information to derive any correlation between the two.

*1.1 Literature Review*

*Internet Symptom Checkers*

Patients have always had to turn to different sources of information when making health decisions. Be it their healthcare provider, friends/family, brochures, or most recently the internet. A 2013 PEW research study found that 35% of US adults had gone online specifically to seek medical information.[6](https://www.zotero.org/google-docs/?8ndinV)

With the prevalence of symptom checkers, many studies in recent years have aimed at analyzing the quality of symptom checkers. One study looking at the accuracy of diagnosis claimed that only in ⅓ of cases was an illness accurately diagnosed in the top 23 symptom checker. This statistic becomes more worrisome when paired with the fact that less than 1/4 of emergent standardized patient evaluations were correctly diagnosed.[7](https://www.zotero.org/google-docs/?12oYHY) Without a medical background, many individuals may have trouble accurately diagnosing their illnesses based purely on symptom checkers rather than medical history and demographic factors.

Aside from the lack of accurately matching symptoms with diagnosis, there is also the danger of misinformation. Online health checkers are not always held up to the same medical rigor or peer-review standards. The internet is full of “fake news” and unconstrained falsities and unfortunately, health information has not escaped this characteristic of the internet. Individuals seeking health information have to shift through inaccurate or unverified medical information that presents as official. With support groups and alternative medicine prevalence, gaining access to valuable and actionable medical advice can become a struggle. Individuals may act on inaccurate information and face health consequences for their decisions based on these sources.[8](https://www.zotero.org/google-docs/?zEc9Hy)

*Patient Behavioral Effect*

Many researchers have also chosen to focus on understanding the makeup of the population of individuals who use the internet as a primary source of health information. Young adults (18-29) were more likely to use the internet than those (60+)—who relied primarily upon their health care providers. In general, women were more likely than men to seek health information whether it be online or from their healthcare provider. Those with a bachelor’s degree were also more likely to obtain health information online than those with less education. Those with higher income were also more likely to get health information primarily from online or their healthcare provider.[6](https://www.zotero.org/google-docs/?J3fR74) The overall trend appears that individuals who use the internet to get health information also trusted their healthcare providers and used the information found online as a supplement to their healthcare provider’s advice rather than a substitute. The main point of digression lay in that those in good health sought out health information online more often than their poor health counterparts.

Previous literature looking at the effect of online health information on the patient-health professional relationship found that patients usually seek health information prior to clinical encounters to better inform visits or to decide whether to seek medical help. Patients also sought online health information after a visit for reassurance or from dissatisfaction with the amount of information provided by the medical provider.[3](https://www.zotero.org/google-docs/?pfvtpG) Internet users were more likely to conduct symptom-based searches for health information looking at only the top results rather than condition-based health searching.[9](https://www.zotero.org/google-docs/?X8cjSC) However, individuals do not always discuss their findings with a medical professional--one study estimated that 59% of people did not discuss health information they found online with their doctor.[10](https://www.zotero.org/google-docs/?biHIIm)

*Negatives / Cyberchondria*

There are many dangers associated with the growth of online health information. Some of the top concerns include patients feeling a false sense of reassurance from symptom checkers and then deciding to delay seeing a doctor or not seek professional help altogether. In a 2013 survey, it was found that 47% of individuals did not confirm their self-diagnosis with a medical health professional.[11](https://www.zotero.org/google-docs/?AH3FX7)

Patients may even try to self-medicate: Dr.Bavishi stated that "If someone were to take the advice of a symptom checker -- which again they're not supposed to -- but if they were to self-treat without the opinion of a physician, the problem is that follow-up may be missed and that can potentially turn into a problem as well."[12](https://www.zotero.org/google-docs/?Qy1yDD)

More prominently among the concerns is the growth of cyberchondria. The sheer quantity of health information online can cause some individuals to become over-anxious about their health.[5](https://www.zotero.org/google-docs/?kCXzYW) Individuals will begin to misdiagnose themselves with serious disorders despite the low likelihood that their symptoms are related to a serious illness. This can also lead to further inefficient use of medical resources with individuals making excessive doctor's appointments and obtaining unnecessary diagnostic testing.[5](https://www.zotero.org/google-docs/?AMqXNB)

*Potential Positives*

While there are many problems associated with individuals relying on online health checkers, there are also many potential benefits that can arise. Having health information readily available and accessible provides an instant resource for individuals struggling to connect with a professional healthcare provider. Numerous popular symptom checkers also have sections informing users when a symptom requires medical attention vs. when over-the-counter medication is more appropriate.[13](https://www.zotero.org/google-docs/?XE9jVA) Using online health symptom checkers effectively can become a powerful tool to reduce inefficiencies within the healthcare industry.

Some patients have begun using the internet to come to doctor appointments with more targeted questions and concerns.[14](https://www.zotero.org/google-docs/?uFG0WF)It also allows doctors to get a better understanding of a patient's symptoms and provide them supplementary resources that they did not cover during an appointment.[14](https://www.zotero.org/google-docs/?9dsElq) Many patients cited having online health information as empowering, giving them more reassurance, and clinical understanding of their condition.[15](https://www.zotero.org/google-docs/?rPorxZ)Overall, patients felt that online health searches had a positive impact on the patient-clinician relationship.[15](https://www.zotero.org/google-docs/?cvnB34) It also allowed patients to play a more active role in their healthcare and make more informed healthy lifestyle choices.[13](https://www.zotero.org/google-docs/?9kOojP)

*1.2 Methods and Methodology*

*Study Sample*

To investigate the correlation between internet usage and the effectiveness/appropriateness of a medical visit, two datasets were used. The first was used to generate a profile on who used the internet for medical searches, and the second was to understand how efficiently people were using medical visits. This project inspected the data from 2003, 2005, and 2007 in its analysis.

The first dataset used in the analysis is the Health Information National Trends Survey, from here on referred to as [HINTS](https://hints.cancer.gov/data/Default.aspx). The purpose of the survey is to access health communication processes, knowledge, attitudes, and behaviors in US adults. Beginning in 2003, almost annually, this survey has randomly sampled a nationally representative group of individuals. For this project, (N=15,869) observationswere used over 2003, 2005, and 2007 to obtain the demographic and socioeconomic variables correlated to the use of online health information.

Using a combination of questions, a variable was derived for accessing whether or not an individual used the internet to search for health information. In the 2003 and 2005 datasets, I used the variable hc24ahealthinfoself which asked respondents “Did you use the internet to seek health information for yourself?” Respondents that responded in the affirmative were coded with a 1 while those who responded “no” were coded as a 0. The participants who answered that they did not go online in an earlier question were not asked this question. However, I recoded the missing value from their response with a 0. In the 2007 dataset, the variable hc02whereseekhealthinfo was used to assess internet usage. It asked respondents what source they used “the most recent time you looked for information about health or medical topic”. Participants were given the options: (1) books (2) brochures, pamphlets, etc. (3) cancer organizations (4) family (5) doctor of healthcare provider (6) internet, etc. I coded those who responded with the internet (58.13% of respondents) as a 1 and the remaining as a 0. This internet usage variable was used alongside demographic variables including sex, age, and race/ethnicity.

The second dataset is the National Ambulatory Medical Care Survey. The [NAMCS](https://www.cdc.gov/nchs/ahcd/index.htm) data, starting from 1993, “collect[s] information on office-based physicians’ adoption and use of electronic health record (EHR) systems, practice information, patient engagement, controlled substances prescribing practices, use of health information exchange (HIE), and documentation and burden associated with medical record systems.” This dataset was used to determine the efficiency of doctor’s appointments. Specifically, the variable “Reason for Visit”, detailing the ICD-10-CM Diagnosis Code, was used to determine the efficiency of a medical appointment. I used (N=31,244) observations from 2003, 2005, and 2007 for my analysis.

*Statistical analysis*

First, from the HINTS data, using the questions specified above, a binary variable was generated to indicate whether an individual (1) used the internet to seek health information or (0) did not use the internet/ uses another source. The three re-coded demographic variables included age (18-29, 30-49, 50-64, 65+ years of age), sex (male, female), and race/ethnicity (non-Hispanic (NH) White, NH Black, Hispanic, NH Asian, Other). I used binary regressions to analyze the relationship between the selected covariates (age, sex, race/ethnicity) and internet usage to seek health information (binary variable). The regression was run separately for the 2003, 2005, and 2007 datasets yielding a unique set of coefficients for each year and formulas to generate a propensity to seek out health information online scores for individuals.

**Table 1: Coefficients of Internet Propensity**

|  | 2003 Coef. | Std. Err. | 2005 Coef. | Std. Err. | 2007 Coef. | Std. Err. |
| --- | --- | --- | --- | --- | --- | --- |
| Age=30-49 | -0.053 | .018 | -0.0193 | .023 | 0.014 | .022 |
| Age=50-64 | -0.156 | .02 | -0.126 | .024 | -0.092 | .022 |
| Age=65+ | -0.394 | .021 | -0.386 | .025 | -0.383 | .023 |
| Sex=Female | 0.087 | .013 | 0.0636 | .014 | 0.058 | .012 |
| Race/Ethnicity=NH Black | -0.112 | .019 | -0.150 | .025 | -0.182 | .021 |
| Race/Ethnicity=Hispanic | -0.233 | .019 | -0.285 | .024 | -0.253 | .021 |
| Race/Ethnicity=Asian | 0.065 | .045 | 0.026 | .05 | -0.143 | .037 |
| Race/Ethnicity=Other | -0.076 | .036 | -0.103 | .036 | -0.099 | .034 |
| Sample Size |  | 5,421 |  | 4,346 |  | 6,102 |

**\*Table 1 lists the coefficients of the logistic regression for each year of the demographic variables to the binary variable of internet use for health information**

The coefficients from the HINTS data binary regression as listed above in Table 1 were applied to the same covariates in the NAMSC data of the respective year to ascertain a variable predicting an individual’s propensity to seek out health information online. Table 2 summarizes the mean generated propensity values within each subcategories by year.

**Table 2: Summary of Propensity to use the Internet for Health Information (Mean)**

|  | 2003 | 2005 | 2007 |
| --- | --- | --- | --- |
| Age |  |  |  |
| 18-29 | 0.543 | 0.599 | 0.703 |
| 30-49 | 0.415 | 0.466 | 0.564 |
| 50-64 | 0.286 | 0.333 | 0.427 |
| 65+ | 0.160 | 0.200 | 0.286 |
| Sex |  |  |  |
| Male | 0.250 | 0.307 | 0.405 |
| Female | 0.355 | 0.392 | 0.480 |
| Race & Ethnicity |  |  |  |
| NH White | 0.325 | 0.373 | 0.464 |
| NH Black | 0.301 | 0.334 | 0.449 |
| Hispanic | 0.253 | 0.279 | 0.393 |
| Asian | 0.157 | 0.208 | 0.306 |
| Other | 0.146 | 0.194 | 0.321 |

**\*Table 2 lists the mean generated propensity to use the internet score for each category by year in the NAMSC sample**

Then, using the NAMSC data a variable ranking the effectiveness of medical care was created using the data for the primary Reason for the Visit. The definition of what is considered an effective medical visit was inspired by the 1986 paper “Use of Medical Care in the Rand Health Insurance Experiment: Diagnosis- and Service-Specific Analyses in a Randomized Controlled Trial” (Lohr, et al.) For the purposes of their research, the paper created medical effectiveness groupings based on symptoms and diagnosis. The grouping of the effectiveness of medical visits in this paper was modeled after these medical effectiveness groupings: "Group 1: Highly Effective Treatment by Medical Care System", "Group 2: Quite Effective Treatment by Medical Care System", "Group 3: Less Effective Treatment by Medical Care System", and "Group 4: Medical Care Rarely Effective or Self-care Effective". Group 1 included conditions such as pneumonia, diabetes, rheumatic disease, etc. Group 2 included asthma, hemorrhoids, hay fever, etc. Group 3 included vertiginous syndromes, cerebrovascular disease, edema, etc. Group 4 included fatigue, fever, throat pain, headaches, etc. The final distribution of effectiveness of treatments accross each year is shown as a percentage below in Table 3.

| **Table 3: Distribution of Effectiveness across Years** |  | | |
| --- | --- | --- | --- |
| Effectiveness of Visit | 2003 | 2005 | 2007 |
| Group 1: Highly Effective Treatment by Medical Care System | 33.12% | 35.97% | 34.66% |
| Group 2: Quite Effective Treatment by Medical Care System | 8.00% | 7.39% | 7.16% |
| Group 3: Less Effective Treatment by Medical Care System | 20.93% | 20.40% | 18.91% |
| Group 4: Medical Care Rarely Effective or Self-care Effective | 37.95% | 36.23% | 39.27% |

**\*Table 3 summarizes the percentage distribution of the effectiveness of appointments in each given year**

To better analyze the relationship between the use of the internet as a source of health information and the effectiveness at which individuals use medical appointments, the effectiveness variable was further broken down into three variables. The first dependent variable is a binary recode of effectiveness: those in group 1 and group 2 are coded as a 1, while observations in groups 3 and 4 are coded as a 0. The second dependent variable ‘highly ineffective’ codes group 4 as a 1 and all other groups as a 0. The third dependent variable ‘highly ineffective codes group 1 as a 1 and all other variables as a 0.

After establishing three dependent variables, we analyzed if there is any correlation between internet use and the efficiency of healthcare visits and any other trends that may emerge in individuals that use the internet for healthcare information. The final analysis was derived using logistic regression to analyze the relationship between the dependent variables measuring effectiveness, the internet use propensity score, and demographic control variables (sex, age, race/ethnicity year).

**Chapter 2. Results**

| **Table 4: Effectiveness Regression** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Coef.** | **SE** | **z** | **p-value** | **90% lower** | **90% upper** |
| **Internet Propensity** | **-2.23\*** | **1.21** | **-1.84** | **0.07** | **-4.22** | **-0.24** |
| **Sex=Female** | **-0.01** | **0.08** | **-0.14** | **0.89** | **-0.15** | **0.13** |
| **Race=NH Black** | **0.17\*\*** | **0.08** | **2.06** | **0.04** | **0.03** | **0.30** |
| **Race=Hispanic** | **-0.27\*** | **0.15** | **-1.82** | **0.07** | **-0.51** | **-0.03** |
| **Race=Asian** | **-0.41\*** | **0.22** | **-1.81** | **0.07** | **-0.77** | **-0.04** |
| **Race=Other** | **-0.55\*** | **0.32** | **-1.75** | **0.08** | **-1.07** | **-0.03** |
| **Age=30-49** | **-0.37\*\*** | **0.17** | **-2.22** | **0.03** | **-0.65** | **-0.10** |
| **Age=50-64** | **-0.45** | **0.33** | **-1.37** | **0.17** | **-0.98** | **0.09** |
| **Age=65+** | **-0.50** | **0.49** | **-1.02** | **0.31** | **-1.30** | **0.31** |
| **year=5** | **0.20\*\*** | **0.06** | **3.17** | **0.00** | **0.09** | **0.30** |
| **year=7** | **0.34\*** | **0.18** | **1.94** | **0.05** | **0.05** | **0.63** |

**\*Table 4:shows the coefficient, standard error, z-score, p-value, and 90% confidence interval of a binary logistical regression between effective treatment and internet use Significance levels: \*p <.1; \*\*p <.05; (N=31,244)**

Table 4 reports the binary logistic regression results of the effective variable and an individual's propensity to use the internet for health information. From the 31,244 observations, the results showed that the coefficient between the effectiveness of medical visits and internet use was -2.226. The 90% confidence interval concluded that the correlation between the two variables is negative. Individuals who use the internet as a source of health information are less inclined to use medical appointments effectively.

Non-Hispanic African American individuals were the only race/ethnicity with a positive coefficient (.167), suggesting they made more effective use of medical appointments. Those aged 30-49 had a coefficient of -.373 with a 95% confidence interval indicating that the younger generations were more likely to use medical appointments ineffectively. The coefficients over years 5 and 7 were positive and rising (.196 and .341 respectively). This demonstrates that over time people are making better-informed decisions and using medical appointments more effectively.

Table 5 and Table 6 detail the results of the same regression run for highly effective and highly ineffective treatment as a binary variable.

| **Table 5: Highly Ineffectiveness Regression** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Coef.** | **SE** | **z** | **p-value** | **90% lower** | **90% upper** |
| **Internet Propensity** | **1.32** | **1.23** | **1.08** | **0.28** | **-0.70** | **3.34** |
| **Sex=Female** | **0.12** | **0.08** | **1.36** | **0.17** | **-0.02** | **0.25** |
| **Race=NH Black** | **-0.14\*** | **0.08** | **-1.71** | **0.09** | **-0.28** | **-0.01** |
| **Race=Hispanic** | **0.15** | **0.15** | **1.01** | **0.31** | **-0.09** | **0.40** |
| **Race=Asian** | **0.44\*** | **0.23** | **1.93** | **0.05** | **0.07** | **0.81** |
| **Race=Other** | **0.39** | **0.32** | **1.21** | **0.23** | **-0.14** | **0.91** |
| **Age=30-49** | **0.11** | **0.17** | **0.65** | **0.52** | **-0.17** | **0.39** |
| **Age=50-64** | **-0.06** | **0.33** | **-0.18** | **0.86** | **-0.60** | **0.49** |
| **Age=65+** | **-0.20** | **0.49** | **-0.40** | **0.69** | **-1.01** | **0.61** |
| **year=5** | **-0.14\*\*** | **0.06** | **-2.14** | **0.03** | **-0.24** | **-0.03** |
| **year=7** | **-0.13** | **0.18** | **-0.71** | **0.48** | **-0.42** | **0.17** |

**\*Table 5:shows the coefficient, standard error, z-score, p-value, and 90% confidence interval of a binary logistical regression between highly ineffective treatment and internet use Significance levels: \*p <.1; \*\*p <.05; (N=31,244)**

| **Table 6: Highly Effective Regression** | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Coef.** | **SE** | **z** | **p-value** | **90% lower** | **90% upper** |
| **Internet Propensity** | **-2.27\*** | **1.28** | **-1.78** | **0.08** | **-4.37** | **-0.17** |
| **Sex=Female** | **-0.10** | **0.09** | **-1.13** | **0.26** | **-0.24** | **0.05** |
| **Race=NH Black** | **0.23\*\*** | **0.09** | **2.66** | **0.01** | **0.09** | **0.37** |
| **Race=Hispanic** | **-0.19** | **0.16** | **-1.25** | **0.21** | **-0.45** | **0.06** |
| **Race=Asian** | **-0.35** | **0.24** | **-1.48** | **0.14** | **-0.74** | **0.04** |
| **Race=Other** | **-0.51** | **0.33** | **-1.54** | **0.12** | **-1.06** | **0.04** |
| **Age=30-49** | **-0.19** | **0.18** | **-1.06** | **0.29** | **-0.48** | **0.10** |
| **Age=50-64** | **-0.14** | **0.34** | **-0.40** | **0.69** | **-0.70** | **0.43** |
| **Age=65+** | **-0.14** | **0.51** | **-0.26** | **0.79** | **-0.98** | **0.71** |
| **year=5** | **0.23\*\*** | **0.06** | **3.60** | **0.00** | **0.13** | **0.34** |
| **year=7** | **0.38\*\*** | **0.19** | **2.08** | **0.04** | **0.08** | **0.69** |

**\*Table 6:shows the coefficient, standard error, z-score, p-value, and 90% confidence interval of a binary logistical regression between highly effective treatment and internet use Significance levels: \*p <.1; \*\*p <.05; (N=31,244)**

The results from highly effective and highly ineffective care reflected similar findings. Individuals with a higher propensity to use the internet were less inclined to have highly effective reasons for medical appointments (coefficient of -2.27). While not as significant, internet users also had a higher coefficient (1.32) when it came to having ineffective medical appointment reasons. As seen before, Non-Hispanic black individuals not only had fewer ineffective appointments (coefficient -.14) but also had a positive correlation with having highly effective reasons for visiting (coefficient .226).

The correlation between highly effective visits and years was also seen growing stronger. In 2005 and 2007, there was a growing positive correlation between having more medically efficient visits (.234 and .385 respectively). The correlation between ineffective visits and the years 2005 and 2007 was also negative (-.135 and -.128 respectively).

**Chapter 3. Discussion**

Contrary to our previous hypothesis, individuals who used the internet as a source of health information did not use medical appointments more efficiently. They not only had more ineffective appointments than their counterparts but more significantly, they had fewer medical appointments concerning symptoms and illnesses where medical care was considered highly effective. This hints that there may be some danger of individuals downplaying the severity of their conditions or feeling that they can benefit from self-treatment due to the internet.

However, there are also some potential biases present in the data and method of analysis. Individuals with a higher propensity to use the internet may be going to the doctor less. Since there is no denominator on which individuals eventually decide to seek medical attention, there exists the possibility of a hidden confounding variable. These individuals might be healthier for some reason, and therefore do not need to make medical appointments for highly effective illnesses.[16](https://www.zotero.org/google-docs/?lsMEIA)

The current predictive model for internet use contains only three demographic variables: age, sex, and race/ethnicity. These limitations were caused by the lack of demographic variables in the NAMSC data. In the future, adding more demographic variables to build the profile of an internet user will likely lead to more accurate results. It will also allow for the addition of more control variables in the final regression analyzing the relationship between one's propensity to use the internet and the effectiveness of their medical visit.

A last emerging trend to draw attention to is the growing efficiency of visits over years. In the years 2005 and 2007, the correlation between the efficiency of visits was positive with a strengthening correlation. Individuals appeared to be making more effective use of medical visits. While this study focused on earlier data to capture the behavior of non-internet users, the quality and quantity of health information have grown over time and may lead to different results in the present.[1](https://www.zotero.org/google-docs/?OpoPgh)

**References**

[1. Nowrouzi B, Gohar B, Nowrouzi-Kia B, Garbaczewska M, Brewster K. An Examination of Scope, Completeness, Credibility, and Readability of Health, Medical, and Nutritional Information on the Internet: A Comparative Study of Wikipedia, WebMD, and the Mayo Clinic Websites. *Can J Diabetes*. 2015;39:S71. doi:10.1016/j.jcjd.2015.01.267](https://www.zotero.org/google-docs/?A22at6)

[2. Bodenheimer T. High and Rising Health Care Costs. Part 1: Seeking an Explanation. *Ann Intern Med*. 2005;142(10):847-854. doi:10.7326/0003-4819-142-10-200505170-00010](https://www.zotero.org/google-docs/?A22at6)

[3. McMullan M. Patients using the Internet to obtain health information: How this affects the patient–health professional relationship. *Patient Educ Couns*. 2006;63(1):24-28. doi:10.1016/j.pec.2005.10.006](https://www.zotero.org/google-docs/?A22at6)

[4. Chong S, Ding D, Byun R, Comino E, Bauman A, Jalaludin B. Lifestyle Changes After a Diagnosis of Type 2 Diabetes. *Diabetes Spectr Publ Am Diabetes Assoc*. 2017;30(1):43-50. doi:10.2337/ds15-0044](https://www.zotero.org/google-docs/?A22at6)

[5. White RW, Horvitz E. Cyberchondria: Studies of the escalation of medical concerns in Web search. *ACM Trans Inf Syst*. 2009;27(4):1-37. doi:10.1145/1629096.1629101](https://www.zotero.org/google-docs/?A22at6)

[6. Smith D. Health care consumer’s use and trust of health information sources. *J Commun Healthc*. 2011;4(3):200-210. doi:10.1179/1753807611Y.0000000010](https://www.zotero.org/google-docs/?A22at6)

[7. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. *BMJ*. 2015;351:h3480. doi:10.1136/bmj.h3480](https://www.zotero.org/google-docs/?A22at6)

[8. Cline RJW, Haynes KM. Consumer health information seeking on the Internet: the state of the art. *Health Educ Res*. 2001;16(6):671-692. doi:10.1093/her/16.6.671](https://www.zotero.org/google-docs/?A22at6)

[9. Mueller J, Jay C, Harper S, Davies A, Vega J, Todd C. Web Use for Symptom Appraisal of Physical Health Conditions: A Systematic Review. *J Med Internet Res*. 2017;19(6):e202. doi:10.2196/jmir.6755](https://www.zotero.org/google-docs/?A22at6)

[10. Diaz JA, Griffith RA, Ng JJ, Reinert SE, Friedmann PD, Moulton AW. Patients’ use of the Internet for medical information. *J Gen Intern Med*. 2002;17(3):180-185. doi:10.1046/j.1525-1497.2002.10603.x](https://www.zotero.org/google-docs/?A22at6)

[11. Kuehn BM. More Than One-Third of US Individuals Use the Internet to Self-diagnose. *JAMA*. 2013;309(8):756-757. doi:10.1001/jama.2013.629](https://www.zotero.org/google-docs/?A22at6)

[12. Ramel D. The doctor is ... YOU! Online medical symptom checkers examined. Computerworld. Published February 12, 2009. Accessed May 11, 2022.](https://www.zotero.org/google-docs/?A22at6) <https://www.computerworld.com/article/2530783/the-doctor-is-----you--online-medical-symptom-checkers-examined.html>

[13. Ryan A, Wilson S. Internet healthcare: do self-diagnosis sites do more harm than good? *Expert Opin Drug Saf*. 2008;7(3):227-229. doi:10.1517/14740338.7.3.227](https://www.zotero.org/google-docs/?A22at6)

[14. Why You Should Never, Ever Self-Diagnosis Using Google. Philadelphia Magazine. Published April 26, 2013. Accessed May 11, 2022.](https://www.zotero.org/google-docs/?A22at6) <https://www.phillymag.com/be-well-philly/2013/04/26/never-ever-self-diagnosis-google/>

[15. Wang J, Ashvetiya T, Quaye E, Parakh K, Martin SS. Online Health Searches and Their Perceived Effects on Patients and Patient-Clinician Relationships: ASystematic Review. *Am J Med*. 2018;131(10):1250.e1-1250.e10. doi:10.1016/j.amjmed.2018.04.019](https://www.zotero.org/google-docs/?A22at6)

[16. Shavers VL. Measurement of socioeconomic status in health disparities research. *J Natl Med Assoc*. 2007;99(9):1013-1023.](https://www.zotero.org/google-docs/?A22at6)

**Acknowledgments**

I would like to express my deepest gratitude to Dean Sherry Glied for her invaluable insight into the methodology of how to answer my question and her timely feedback and comments during the completion of this thesis. I would also like to extend my sincere appreciation to Bonnie Lawerence at NYU Library’s Data Service for helping me navigate my datasets. Lastly, I am very grateful to all the faculty members in the Economics Honors Program at New York University Shanghai for their support and extensive knowledge.