ABSTRACT

ZIELINSKA, OLGA ANNA. What are the Chances? Examining the Effect of Displaying Probability Rates on Web Health Searches. (Under the direction of Dr. Christopher B. Mayhorn).

The Internet offers many benefits to people seeking health information, such as the convenience of accessing information at any time, and the protection of viewing information anonymously; however, such information is unregulated and can be misinterpreted. Escalation, the observed increase in medical severity of search terms within a single search session, could occur. For example, escalation occurs when an initial search for "headache" leads to a later search for "brain tumor". Researchers have recommended including incidence rates to reduce escalation; however, this phenomenon has yet to be tested empirically. One-hundred-and-fifty undergraduates were randomly assigned to one of three presentation groups (control, pictorial, and numeric) where they evaluated four search results pages. Incidence rates were not displayed in the control whereas participants in the pictorial condition saw incidence rates displayed as bar graphs and those in the numeric condition saw incidence rates displayed as percentages. Escalation was evaluated using the susceptibility and severity measures from the Risk Behavior Diagnosis Scale. Four symptoms were evaluated and each symptom was paired with four conditions: two benign and two serious. Perceived severity increased from pre-search to post-search ($p<.001$), but there were no display group differences. Benign susceptibility was highest when shown numerically, followed by pictorially, and lowest in the control. The numeric group was significantly higher than the control group ($p=.001$) and the pictorial group ($p=.037$); however, there were no differences between the control and pictorial group ($p<.10$). Similarly in the serious condition, the highest ratings were in the numeric group, followed by pictorial and the lowest

susceptibility in the control. Numeric was significantly higher than the control (p=.003) and pictorial (p=.028), but there was no difference between the pictorial group and the control (p<.10). Although susceptibility was higher when incidence rates were present for both benign and serious conditions, rates were higher for benign conditions than the serious conditions (p<.001) suggesting that people are not escalating. The findings from this study could be helpful in understanding how patients comprehend healthcare information and could conceivably provide direction for how health care professionals distribute information to their patients.

.

What are the Chances?
Examining the Effect of Displaying Probability Rates on Web Health Searches

by
Olga Anna Zielinska

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Psychology

Raleigh, North Carolina

2015

APPROVED BY:

Dr. Christopher B. Mayhorn Committee Chair	Dr. Anne McLaughlin
Dr. Jing Feng	

BIOGRAPHY

Olga Anna Zielinska is a graduate student and research assistant in the Human Factors and Applied Cognition Program. Prior to moving to Raleigh, NC, She received a Bachelor of Science in Biomedical Engineering with a specialty in Neuroengineering and a minor in Psychology from Drexel University. During her undergraduate career, Olga worked as an engineering assistant in a consulting firm, where she discovered the field of human factors. The intersection of human behavior and system interaction captured her interest and she decided to pursue a doctorate to conduct research in the field of human factors.

ACKNOWLEDGMENTS

First, I would like to thank my advisor, Dr. Chris Mayhorn, for his continuing support and guidance. From the first day of graduate school, he involved me in multiple research opportunities that exposed me to a variety of human factors specialties and allowed me to explore my individual research interests. Thank you to my committee, Dr. Anne McLaughlin and Dr. Jing Feng, for their insight and thought-provoking questions, especially during the idea formation and proposal process.

I would also like to thank my research assistants, Rebecca McNulty and Edmund Buckman, for their hard work and dedication in facilitating the data collection phase of this project. It would not have been possible without you.

Finally, I would like to thank my parents for their continual encouragement and support during this process. Their sacrifices and unyielding work ethic have inspired me to make the most of my education and the opportunities afforded to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Internet is a quick, easy, and free way to access information. One topic that is largely searched on the Internet is health care because of the unlimited amount of information that can be searched at any time of day anonymously, reducing the chance of judgment on sensitive subjects; however, the health care information available online is largely unregulated leading to some inaccuracies and misinterpretation (Raine et al., 2000; Starcevic & Berle, 2013). For example, when searching for a medical symptom, such as a headache, brain tumor could be thought of as the most likely condition because it is at the top of a search results page; while, it could simply be at the top of the list due to the fact that a lot Internet users were curious about the topic and clicked on it more often than a more benign condition, such as tension headache. This trend could lead certain health anxious users to feel overwhelmed and frightened about what they find online. Multiple investigators have made recommendations on how to improve the search result displays of medical searches with one common theme: adding in incidence rates of illnesses (Aiken et al., 2012; Starcevic & Berle, 2013; White & Horvitz, 2009). By listing the probability of actually having a disease, researchers believe this will reduce anxiety in Internet users; however, there is a lack of research to support the theory. The purpose of the current research is to investigate the effects of adding probability rates to Internet health search results.

Problem Statement

A survey done by the Pew Research Group in September 2012 found that 81% of adults in the United States use the Internet. Seventy-two percent of those Internet users have looked online for health information in the last year. From the users that have looked for online

health information, 59% have gone online to specifically find out what medical condition they have (or think they have) or someone else has. This represents over one third of the adults in the United States (Fox & Duggan, 2013). In a separate report, Google stated that health-related websites rank among the top 1000 websites with about 118 million unique monthly visitors. WebMD has also reported having 112 million unique monthly visitors to their website. (Lauckner & Hsieh, 2013).

The Internet can offer many benefits to people seeking information on their health. Users have cited the convenience of accessing the information at any time, the protection of viewing information anonymously and without judgment, as well as being able to access more information online than from any other source as the top three reasons they go online for health information (Raine et al., 2000). Additionally, the information can bring answers fast with no cost (Starcevic & Berle, 2013).

Users can also feel empowered with the information they find online and become more comfortable bringing up health care issues to their physician (Lauckner & Hsieh, 2013; Starcevic & Berle, 2013). For example, Seckin (2010) found that cancer patients used the Internet to increase their knowledge about cancer, learn about different treatment options, as well as the advancements in patient care. However, users need to be cautious about the information they find online, because most is unregulated. Without the proper training to navigate through the search results, users may misinterpret their findings. For example, some users may believe that medical conditions listed at the top of a search results page are the best explanations for their symptoms, regardless of the true incidence rate. As a result, this could lead to increased anxiety as depicted in an excerpt from an article seen below:

It had started with tingling and numbness in her legs. For months she had searched the web in an attempt to find out what kind of disease she had. She ended up in a neurological chat room, and came to a devastating conclusion - she must have multiple sclerosis.

"I was reading what other sufferers had posted there," she says now. "It sounded exactly like what I had. I was sure I was dying."

Indeed, Melissa Woyechowsky was ill. Yet she was not manifesting the early symptoms of MS (exacerbated, she later came to think, by Lou Gehrig's disease and several other self-diagnosed maladies).

Rather, she was suffering from something that was already far more advanced. At the age of 30 she was one of the world's first out-of-control cyberchondriacs.

If hypochondria is the excessive fear of illness, then cyberchondria is that fuelled by the huge amount of medical information that is now available on the Internet. (Valley, 2001).

Although this behavior seems extreme, studies have shown that escalation can occur in Internet health studies. White and Horvitz (2009) have defined escalation as the change in a web search from basic symptoms to reviewing of information on serious illness. Recommendations have been made in the literature to list the probabilities or incidence rates in the search results of each illness to reduce escalation and anxiety in users.

Purpose of Research

Although recommendations have been made to add incidence rates to illness listings during health web searches, few websites have adopted the practice. This may be due to the

fact that there has been little research testing the effects of adding incidence rates to medical web searches. The purpose of this research is to learn more about the effects of escalation with the display of probability rates next to the illnesses during health Internet searches and if such an intervention will minimize Internet users' escalation and anxiety while performing their searches of symptoms on the Internet.

Literature Review

Knowledge Base

Prior to empirical literature review, it would be beneficial to define key terminology that is used throughout the remainder of the proposal. Specifically, the eight following terms will be explained: escalation, health anxiety, hypochondria, cyberchondria, severity, susceptibility, health literacy, and incidence rate.

Escalation is defined as observed increases in the medical severity of search terms used within a single health-related search session (White & Horvitz, 2010.) For example, escalation would be used to describe someone who searches "headache" online and then searches "brain tumor". With that same example, if someone searched headache and then searched tension headache or migraine, they would not be considered to display escalation. Health anxiety is described as the fear, belief, or misinterpretation of bodily signs and symptoms to indicate the presence of a serious illness. Hypochondria is health anxiety in its most severe form (Aiken et al., 2012; Baumgartner & Hartmann, 2011). Cyberchondria is health anxiety exacerbated by the exposure to Internet based information. It is also referred to as the excessive or repeated search for health information on the Internet, driven by distress or anxiety about health, which then amplifies the same distress or anxiety (Starcevic & Berle,

2013). Severity is defined as the magnitude of harm expected from a threat, while susceptibility is the likelihood that a specific person will experience a threat (Witte, Meyer, & Martell, 2001). Health literacy is the ability to perform basic reading and numerical tasks required to function in a health care environment (Lauckner & Hsieh, 2013). Finally, incidence rates, also referred to as likelihood, is the prevalence or the probability of being diagnosed with a certain condition. One example of an incidence rate is that 1 in 10,000 people are diagnosed with a brain tumor when complaining of a headache (White & Horvitz, 2009).

Empirical Research

The literature review focuses on two aspects involved in the research of searching for health information online. The first section examines escalation through the perspective of web search logs. Computer algorithms analyzed users' behaviors through the use of search engines. The second and more elaborate aspect focuses on the psychological effects of escalation including, but not limiting to, the perceived severity, perceived susceptibility, and perceived anxiety of searching for health symptoms online.

Escalation through Computer Search Logs

White and Horvitz (2009) examined hundreds of thousands of Internet Explorer users' search logs over an 11 month period that consented to having Windows Live Toolbar. Windows Live Toolbar is a plug-in for Internet Explorer used to help improve performance of the Internet experience. The logs had recorded searches through all major search engines, such as Google and Yahoo. Two percent of the log queries were health related and 250 thousand users did at least one medical search during the 11 month period. They further

narrowed down the area of interest to queries that contained one of 12 medical symptoms (breathlessness, chest pain, dizziness, fatigue, fever, headache, insomnia, lump, nausea, rash, stomach pain, and twitching) leading to 8732 subjects and 11,158 sessions. Researchers found that 5.3% users escalated, 7.4% of users did not escalate, and 87.3% had no change. A computer algorithm was used in the study to indicate which search combinations were escalations and non-escalations. The search results that were not specifically defined in the algorithm as an escalation or non-escalation were placed into a no change category. The high number of no change was due to the fact that either the computer algorithms did not cover the breadth of symptoms and conditions or users typed in the symptom in the same search as the illness or searched the illness first and then the symptom.

In a similar web search log analysis study in 2010, White and Horvitz found that 69% of people escalated when serious conditions were presented first on the search results page. The percentage of escalation dropped to 33% when the benign conditions were presented first on the search results page. Presentation order also had an effect on the perception of severity, susceptibility, and anxiety in subsequent studies.

Psychological Surveys

White and Horvitz included a survey in their 2009 study that was distributed to 5,000 randomly chosen Microsoft Employees, of which 515 volunteers completed the survey. One requirement for completing the survey was that employees performed at least one search for health related information. Although White and Horvitz have previously focused on analyzing search domain results with computer algorithms, it is important to note that the employees that completed the surveys did not have their search histories analyzed. The

responses were self-report. The authors stated that although Microsoft employees were not necessarily representative of the population of online users, there was no evidence that their online medical search experiences were significantly different from the population. The survey revealed that 40% of users reported being concerned about having a serious medical condition based on an Internet search when no condition was diagnosed. About 90% of people admitted to at least one instance where searching the Web for a symptom of a basic medical condition led to reviewing information of a more serious illness. Additionally, 78% of people at one time considered the order of the web search results as representing the likelihood of illness with the more likely presented on top (White & Horvitz, 2009).

Rare and serious conditions may attract more attention from worried users who believe they are susceptible to the condition, as well as curious users who are fascinated about the information that they found online, regardless of health anxiety (Starcevic & Berle, 2013). This in turn raises their ranking in the web search results based on popularity. If people interpret those results as likelihood rankings then it could reinforce their anxiety and fears, when more serious conditions are at the top of the list, even when they are the least likely result of the symptom. This is a vicious cycle making people more anxious and confirming their fears with the web search.

Lauckner and Hsieh hypothesized that Internet users take advantage of anchoring and availability heuristics when making judgments of health care search results. Specifically, for the anchoring heuristic, Internet users look at the first search result and compare all subsequent search results to the initial result. Typically judgment biases ensue and users tend to favor the first result they saw. Therefore, when searching for a symptom, if the most

serious results are at the top, then users are more convinced that the symptom is caused by a serious condition. When examining the availability heuristic in the health web search realm, judgments are made with how frequently a diagnosis is mentioned. For example, cancer is often mentioned in the media to raise awareness and promote the importance of screenings, therefore, cancer receives a lot more attention than a more benign condition making users more familiar with the condition. Lauckner and Hsieh have predicted that the more frequently a condition appears in a search result the more likely that they will believe they have the condition. Four experimental conditions were created: serious illnesses were presented either frequently and in the beginning, frequently and at the end, sparsely and in the beginning, or sparsely and at the end. Participants were asked to rate their severity and susceptibility of several symptoms on a 7 point Likert Scale.

Lauckner and Hsieh (2013) found that there was a significant effect of the placement on severity, with serious illnesses placed first yielding higher perceived severity. Researchers also found that the more frequently a disease appeared in the search results, the higher the perceived susceptibility. Additionally, the higher the perceived severity of an illness as well as the higher the perceived susceptibility of the disease, the more the participants felt frightened and overwhelmed. Interestingly, individuals with lower health literacy had a higher perceived severity of an illness. Authors mentioned that individuals that are more literate in health matters, may have contextual knowledge regarding the probability of having a serious disease given a symptom.

Anxiety levels can also have an effect on online health searches. Baumgarnter and Hartmann (2011) hypothesized that health anxiety is related to online health searches. The

authors measured health anxiety using the Whitely Index, a 14 item self-report inventory that measures hypochondria. Researchers used a 1-4 scale instead of the traditional dichotomous option and reported a Cronbach Alpha of 0.89. They also asked participants to rate how they felt after their last online health information search, indicating whether they were overwhelmed, confused, frustrated, frightened, relieved, or reassured by the information they found online. Through a multiple regression analysis they found that health anxiety was predictive of feeling frightened, overwhelmed, and confused by the information users found online. No relation was found between health anxiety and being relieved or reassured.

Baumgartner and Hartmann then asked participants to look up a specific disease, while rating a set of negative responses. Results indicated that health anxiety was a significant positive predictor of believing they were already infected with the disease, the perceived risk of being infected in the future, and overall worrying about the disease. Although Baumgartner and Hartmann used different terminology, this study had similar findings to the Lauckner and Hsieh study, that web searches can increase the perceived severity, perceived susceptibility, and anxiety of Internet users, especially if people are health anxious.

Finally, Muse et al. (2012) conducted a retrospective study where they compared 2 groups of health anxious people (high anxiety vs low anxiety) with regard to their Internet search habits. Health Anxiety was measured by the Short Health Anxiety Inventory (SHAI), an 18 item self-report measure with a Cronbach Alpha of .90. To ensure the groups were different from one another, the researchers measured the health anxiety of over 200 participants, but only included individuals from the top and bottom quartiles for the study. To measure Internet habits, they asked participants if they used the Internet as a source of health

information, how often they searched for health information, how long the sessions lasted when they searched for health information, how distressing they found the health information they found online, what kind of information they were searching for online, what sources they used to find the information, as well as how accurate they perceived the health information they found online to be. Health Anxiety scores were found to be predictive of how frequently users searched for health information and the level of distress they felt. They found that 17% of those with low health anxiety used the Internet to find health information once a month, while 74% of the high anxiety group used the Internet to search for health information at least once a month. High Anxiety individuals also spent significantly more time searching for health information than those in the low anxiety group.

Although each of these articles provides insight into human behavior in online health searches, they do have some limitations. White and Horvitz looked at trends in big data of Internet Explorer users' search logs. This study does have high external validity because Internet Explorer was the Internet browser used by the majority of Internet users in late 2008 and early 2009 (~67%); however, there was little control by the experimenters significantly lowering the internal validity value of the results, making it difficult to conclude that the Internet searches do in fact increase perceived severity, perceived susceptibility, and anxiety in Internet users (Statcounter, 2014). The Internet users could be escalating due to the fact that they are curious about the results. White and Horvitz did conduct a survey of employees to capture their self-reported Internet search habits and search result perceptions, but only a small percentage of employees participated, questioning the motivation of the employees who did participate. Also the employees may not have been completely honest for fear that

their employer may see the results and it could affect their job status. Additionally, the surveys were never correlated with their actual search history. Stating that there is no evidence proving that Microsoft Employees are different than typical Internet users is not as statistically valid as collecting a survey from Internet users and proving it statistically.

Lauckner and Hsieh (2013) had a much higher level of internal validity because they placed participants in one of four experimental conditions; however, they only provided one Google Search page with 10 active links. The ten links led to websites created by the researchers, with information paraphrased from trusted websites. Although they maintained experimental control, they did lose some external validity, because this may not be representative of what users would experience in the real world when searching symptoms through Google. Similar issues are seen in the study by Baumgartner and Hartmann (2011) where they manipulated a Google search results page for participants. Muse et al. (2012) had trouble with both external and internal validity for their study. Researchers based the majority of their conclusions on a retrospective self-report. Self-report measures can be subject to bias due to social desirability. Also self-reports may not be necessarily representative of how they are actually feeling or behaving. When participants have to think back to prior experience, new variables are introduced, such as the participant forgetting details or forgetting how they truly feel due to an extended period of time since their last search. Although each study may have its shortcomings, collectively they do provide similar results and theories on human behavior during Internet health searches. They each also provide recommendations on how to reduce anxiety, severity, and susceptibility during searches.

Starcevic and Berle suggested educating users to become more health literate. This will allow the users to make more educated decisions when conducting Internet health research. White and Horvitz recommended minimizing the amount of times serious conditions appear in the search results. This may lower users' perception that a serious condition is likely to happen and make him or her feel less susceptible to the condition. Several researchers (Aiken et al., 2012, Starcevic & Berle, 2013, White & Horvitz, 2009) proposed that the addition of a disease's incidence rates will help reduce base rate neglect. Base rate neglect is failing to consider the prior probabilities of an event. By taking into consideration the incidence rates, users are predicted to feel less anxious about the disease and have a lower likelihood of escalating; however, do users actually understand incidence rates?

Incidence rates or risks of being diagnosed with a condition are usually presented in one of three formats: percentages (e.g., 20%), proportions (e.g., .20), or frequencies (20 in 100; Ancker & Kaufman, 2007). Lipkus, Samsa, and Rimer (2001) discovered that people more accurately could assess risk when it was presented as a percentage (85% correct) than if it was presented as a frequency (78% correct) or proportion (49% correct), but it is important to note that the majority of the sample were highly educated women thereby making it difficult to generalize the results to the public. A similar study by Grimes and Snively (1999) compared patients' abilities to interpret frequency rates versus proportions (2.6 per 1000 and 8.9 per 1000 vs 1 in 384 and 1 in 112, respectively) of Trisonomy 21 risks. Women, the only gender included in this study, more accurately identified the more frequent risk when the incidence rates were displayed as frequency rates (73%), as compared to proportions (56%). A more demographically expansive study by Waters et al. (2006) compared different

numerical formats (frequency and percentages) and examined the possible benefits of bar graphs as compared to text when examining the benefits and risks of taking a cancer medication. Researchers found that participants were significantly more accurate with percentages (67.4%) than with frequencies (63.4%, $p<.01$). Additionally, participants were more accurate when they had bar graphs (67.5%) than when they were presented only text (63.4%, $p<.01$). No statistical comparisons were given between numerical formats and bar graphs; however, percentages and bar graphs only differed by a .1% accuracy rate. Participants in this study also indicated their preferred methods of receiving risk information as "words" (39%), "pictures" (23%), "no preference" (20%), and "numbers" (18%) suggesting bar graphs are more approachable than numbers are; however, it is important to note that accuracy rates for each of the mediums were in reverse order of preference (numbers: 73%, no preference: 67%, pictures: 63%, and words: 62%). Trends in these studies suggest that incidence rates are understood most when presented as percentages with bar graphs as a close comparison; however, participants would prefer bar graphs over numbers. Another concerning trend is that accuracies in understanding risk rates peaked at 85% with an average close to 70%, meaning that on average 30% of participants do not fully understand accuracy rates. Will the addition of incidence rates actually help lower anxiety or add to cognitive load that users have when searching for medical advice online. Cognitive load refers to the cognitive resources used during learning and problem solving (Chandler & Sweller, 1991). When irrelevant information is presented to users, it could increase cognitive load unnecessarily and impede their skills of processing and interpreting the situation at

hand, or in this case impede their ability to decide which illness is most probable for their symptoms based on the search results.

Additionally, it is important to note that in previous decision making research, providing incidence rates was assumed to make an impact on perceived consequences; however, when there were strong consequences, such as cancer or winning the lottery, people became insensitive to probabilities, instead day dreaming about the possibilities of the consequences, such as what they would buy when they won the lottery, no matter how low the chances of actually winning (Slovic et al., 2005). This research suggests that the incidence rates that are recommended to reduce users' anxieties may not be effective when life-threatening conditions, such as cancer, are presented, where users could imagine the horrible consequences of having the disease, instead of considering the small likelihood that they have the disease.

Finally, previous education could have an impact on the perception of incidence rates. Waters et al. (2006) have found that participants had varied levels of experience accurately identifying risk in percentages, frequencies, bar graphs, or text, depending on education. Specifically, high school education or less had the lowest average accuracy rate (48%), followed by some college or 2 year degree (62%), and those who graduated with a 4 year college degree (75%). Additionally, numeracy, a mathematical literacy level, has been associated with being able to interpret incidence rate information more accurately, with higher numeracy associated with higher accuracy in interpretation of incidence rates (Hawley et al., 2008). In the same study, individuals of all numeracy levels were able recall risk information about a medication the best when it was presented in either a bar graph form or a

table that listed frequencies than when presented in a pie chart, pictograph, sparkplug, or clock mode; however in a separate study, participants who are low in general literacy have also been found to have a better comprehension and recall of health materials when they were presented in icons or illustrations (Wilson & Wolf, 2009), suggesting that graphs may be a more approachable and understandable medium to present incidence rates.

Research Questions and Hypotheses

Recommendations have been made by a variety of sources with the previous literature that the addition of incidence rates of illnesses in Internet search results should decrease anxiety levels and escalation occurrences; however, previous incidence rate research suggests that on average 30% of users do not understand incidence rates. Additionally, decision making research suggests that people may become insensitive to incidence rates if they are presented with sharp consequences such as cancer.

Incidence rates can be presented in a multitude of ways, such as percentages, frequencies, proportions, and bar graphs to name a few. Pictorial forms have been recommended to people of all literacy levels (both health and mathematical) to make it easier to understand; however, will this make an impact on the anxiety levels, perceived severity, and perceived susceptibility levels? Additionally, a higher health literacy has been correlated with lower anxiety levels in Internet health searches; will this theory hold when incidence rates are present, or will lower literacy health individuals have reduced anxiety with the presence of incidence rates? Finally, there are individuals who are predisposed to feeling more anxious about health matters, in its most extreme form, known as hypochondria. Does hypochondria

impact the relationship between health web searches and anxiety, perceived severity and perceived susceptibility?

In this study, the following research questions are addressed, along with the following hypotheses:

RQ1: Will displaying the incidence rates of each illness change the severity, susceptibility, and anxiety levels?

Hypothesis 1:

H_0: Severity, susceptibility, and anxiety levels will be the same with and without the presence of incidence rates.

H_1: Severity, susceptibility, and anxiety levels will be lower with the presence of incidence rates than with no display of incidence rates.

RQ2: Does the display type of the probability rate (pictorial versus numeric) change the severity, susceptibility, and anxiety levels?

Hypothesis 2:

H_0: Severity, susceptibility, and anxiety levels will be the same whether the probability rate is displayed as a number or pictorially.

H_1: Severity, susceptibility, and anxiety levels will be higher when displayed as a number than when it is displayed pictorially.

RQ3: Does hypochondria moderate the change in severity and anxiety levels?

Hypothesis 3:

H_0: Severity and anxiety levels will not be moderated by hypochondria levels.

H_1: Severity and anxiety levels will be moderated with the level of hypochondria. Specifically, participants with low hypochondria will not have a change in severity and anxiety levels after a health web search, while, participants with high hypochondria will have a change in severity and anxiety levels.

RQ4: Does the literacy level of a person (health and numeric) moderate the change in severity and anxiety levels?

Hypothesis 4:

H_0: Severity and anxiety levels will be not be moderated by literacy levels.

H_1: Severity and anxiety levels will be moderated by literacy levels. Specifically, participants with low health and numeric literacy will have a change in severity and anxiety levels, while, participants with high health and numeric literacy will not have a change in severity and anxiety levels.

These research questions and hypotheses will be tested using the methodology described in the next section.

METHOD

Restatement of Major or General Hypothesis

The objective of the study is to determine if users' perceptions of a symptom will change during an Internet health search depending on the presentation of incidence rates. Previous studies have found that presentation order of illnesses in the search results for health symptoms can increase anxiety, severity, and susceptibility in Internet users. Recommendations from previous studies have suggested that adding incidence rates of actually having the disease could help to reduce the anxiety; however, this theory has not

been tested. Studies have also found that a user's level of hypochondria and health literacy could have an effect on his or her perceived anxiety, severity, and susceptibility. This study explores if displaying the incidence rates does change the severity, susceptibility, and anxiety levels. Additionally, efforts are made to determine if the type of incidence rate display changes the severity, susceptibility and anxiety levels. Finally, the study examines if hypochondria levels and literacy levels, specifically health literacy and numerical literacy, moderate anxiety and severity.

Population, Subjects, and Sampling

The purpose of this study is to observe the behavior of a sample of the Internet User Population that researches symptoms of a potential illnesses through a medical search engine, specifically WebMD.

WebMD, a popular medical search engine website, currently shows a bar graph that displays the probability of having a disease for each symptom reported. This website was downloaded and presented as a local copy to participants. WebMD was manipulated for each experimental condition. The three experimental conditions included in this study are: control, numeric, and pictorial. The control group had the area where the bar graph is presented removed. The numeric group had an incidence rate displayed as a percent in place of the bar graph. The pictorial group had the current layout of the WebMD website with the incidence rate displayed as a bar graph.

A convenience sampling of undergraduate psychology students of North Carolina State University were recruited through the Experimetrix Website. Students received class credit for their participation in this study.

A power calculation was performed to estimate the number of participants necessary. As explained further in the data analysis section, a 3 group repeated measures MANOVA was performed. Unfortunately, previous research studies did not reveal their effect sizes for reference. Cohen's F recommended effect sizes are large (.40), medium (.25), and small (.10; Faul et al., 2007). If there was a large effect size (.40), with an alpha level of .05 and power value of .80, a sample size of 64 would be necessary. To be conservative, a medium effect size of (.25) with similar parameters would yield a necessary sample size of 158. For purposes of time and available resources, a sample size of 150 people was used for this study and this yielded an ability to detect an effect size of .26.

Research Design

A 3 x 2 mixed factorial design was implemented in this experiment. The following independent variables were manipulated across trials; condition (control, numeric, and pictorial) and time (pre-search and post-search). Condition was compared between subjects, while time was assessed as a within subject measure.

Measurement

Measurements for this study are divided into two main sections: demographic and dependent variables. The demographic section is further divided into age, sex, race, education level, major, previous medical experiences, working memory, hypochondria, health literacy, and numerical literacy.

Previous medical experiences assessed familiarity with medical conditions. A list of approximately 50 conditions was given to the participant and they were asked to rate them on a scale of 1 to 7 that used the following anchors:"1" indicated that the participant has never

heard of the disease; “3” indicated they have heard the of disease, but they do not know of someone who has had it; “5” indicated they knew someone who has had the disease, but were not close with them; and “7” represented that the participant or a close family member/friend has previously or currently has the disease. Additionally, survey questions inquired about participants’ previous exposure to medicine, such as participant’s mother’s occupation, father’s occupation, and previous work experiences. The participants were also asked about previous habits of searching the Internet, including if they have searched for medical symptoms in the past online and if so how often they used the Internet to search for medical symptoms. Working memory was assessed using the operation span test presented via an E-prime software program from The Georgia Tech School of Psychology Attention and Working Memory Lab. This test measured the memory abilities and attention abilities of each participant. This test has shown a test-retest reliability of .83 (Redick et al., 2012.)

Hypochondria was measured by the Whitely Index, a fourteen item scaled measured on a 7 point Likert Scale. A response of “1” indicated the participants strongly disagree and “7” indicated the participants strongly agree with presented statements. Previous research has used the Whitely Index on a dichotomous scale. To increase variance and keep uniformity in comparison to other scales being used in the current study, the Whitely Scale was altered to include a 7 point Likert scale. The dichotomous version has shown a high test-retest reliability (approximately 0.90, Salkovskis et al., 2002) and high validity. This scale was developed and used on patients diagnosed with hypochondria and those which were not. Whitely Index questions include: Do you often worry about the possibility that you have got

a serious illness; and if a disease is brought to your attention (through the radio, television, newspapers or someone you know) do you worry about getting it yourself?

Health Literacy was measured through the Rapid Estimate of Adult Literacy in Medicine (REALM). This is a 66 item questionnaire where respondents were asked to read words aloud to measure reading level. Words included terms such as medication, occupation, irritation, and inflammatory, among many others. Participants able to correctly read 61 through 66 words were deemed to have a high school reading level and able to read most patient education materials. If patients could correctly read 45 to 60 words, they are at a 7th to 8th grade reading level and might struggle with most patient education materials. Previous research suggests that the REALM has a high reliability and validity (Parker et al., 1995). Finally, mathematical literacy, also known as numeracy was measured using the Berlin Numeracy Test. Participants answered four mathematical questions such as: Imagine we are throwing a five-sided die 50 times. On average, out of these 50 throws how many times would this five-sided die show an odd number (1, 3, 5)? Responses were scored such that participants were given 1 point for each correct answer they provided. From previous research, it is known that the Berlin Numeracy Test has a reliability coefficient of .59 (Cokely et al., 2012). Additionally, numeracy items specifically related to medical situations from the Lipkus et al. (2001) were included. These items have been shown to have a high reliability ($\alpha = .78$).

The second set of measurements are the dependent variables that were tested. This group includes severity, susceptibility, and anxiety. Each of these variables were evaluated on a 7 point Likert Scale with the following anchors: “1” indicating that the participant strongly

disagrees and “7” indicating that the participant strongly agrees. Severity was defined as the magnitude of harm expected from a threat or the significance or seriousness of a threat (Witte, Meyer, & Martell, 2001). The Risk Behavior Diagnosis Scale evaluated severity in three questions: (Symptom) is a serious threat; (Symptom) is harmful; and (Symptom) is a severe threat. These questions were used in the Lauckner and Hsieh (2013) study and have shown a high reliability rating ($\alpha = .93$). Lauckner and Hsieh also used the susceptibility measures in their study with a reliability rating of .76. Susceptibility is the likelihood that a specific person will experience a threat, the degree of vulnerability, personal relevance, or risk of experience a threat. The questions were taken from the risk behavior diagnosis scale and adapted to fit their study. In the current study, three questions were used to assess susceptibility: If I have (symptom), I am at risk for having (serious condition); It is likely that I have (serious condition) if I experience (symptom); (Symptom) is nothing to worry about. Finally, anxiety was measured via self-reports of how overwhelmed or frightened an individual felt when presented with the symptoms. Anxiety was measured with the six-item short form of the Spielberger State-Trait Anxiety Inventory (STAI). Participants rated their agreement on a Likert scale from 1 to 7 for the following statements: I feel calm, I am tense, I feel upset, I am relaxed, I feel content, I am worried. A rating of “1” indicated that they strongly disagree and a “7” indicated that they strongly agree with the presented statements. Previous studies have tested the six item inventory with a reliability coefficient of .82, when participants rated their agreement on a scale of 1 to 4 (Marteau & Bekker, 1992.) The STAI inventory was adjusted for the current study such that it used a 7 point Likert scale to be consistent with the remaining scales.

Procedure

There were three experimental conditions for the present study. The first displayed incidence rates as bar graphs, the second displayed incidence rates as percentages, and the third condition is the control with no incidence rate display.

Assignment of experimental condition was predetermined before the start of the study by a random number generator with 50 people per condition. The order was on a sheet of paper to assist the research assistants to properly assign participants to conditions.

Prior to coming into the lab, participants completed a Qualtrics survey emailed to them with the consent form. Upon agreement participants provided their age, sex, race, education level, major, and previous medical experiences. Finally participants answered the 14 item questionnaire evaluating their level of hypochondria.

Once participants arrived to the lab, they were brought into the office, seated in front of the computer, and asked to complete the cognitive battery that evaluated working memory. The research assistant administered the REALM test for Health Literacy and the extended Berlin Numeracy test.

A vignette in the form of a story was presented to the participant stating that they have had a medical symptom for a few days. Participants were asked to give an initial diagnosis and rate their initial severity and their initial anxiety level. The participants then performed a WebMD search on the symptom. There was no time limit on the search. Once the participant was satisfied with the search results, they were asked to give a final diagnosis and rate their severity, their anxiety, and their susceptibility. Note that susceptibility was only assessed after the web search to reduce undue influence of mentioning specific diseases prior

to the health web search. The participants were allowed to reference the WebMD website while completing the final diagnosis, severity, susceptibility, and anxiety. This search process was repeated for four symptoms.

The four symptoms that were tested included chest pain, headache, abdominal pain, and fatigue. These symptoms were presented in a random order to minimize order effects. For the susceptibility measures, four specific medical conditions were listed for each of the four symptoms (2 benign and 2 serious). The combinations of symptom and severity of consequences are illustrated in Table 1 below.

The WebMD website was downloaded to the local desktop and hosted from there to imitate a realistic search experience. This process helped to control any website updates or changes that might have occurred during the course of the experiment, thereby influencing the validity of the results. Once the participants completed diagnosis for each of the four symptoms, they were debriefed and thanked for their participation.

Table 1. Four symptoms and associated consequences to be used for the susceptibility measures

Symptom	**Benign Condition**	**Serious Illness**
Chest Pain	Panic Attack Muscle Strain	Heart Attack Pulmonary Embolism
Headache	Migraine Tension Headache	Brain Tumor Meningitis
Abdominal Pain	Food Poisoning Constipation	Appendicitis Pancreatic Cancer
Fatigue	Anemia Low Blood Sugar	Chronic Kidney Disease Type 2 Diabetes

RESULTS

The results section contains four sections. The first section provides descriptive statistics of the demographic and dependent variables. The second section reviews the effect of each demographic variable on each post-search dependent variable through a multiple regression. The third section examines the first two research questions posed earlier regarding incidence rate presence (present versus not present) and display type (control versus numeric versus pictorial) differences on perceived severity, perceived anxiety, and susceptibility. Two statistical analyses were done for both questions. First, a repeated measures MANOVA examined if there were significant differences in the change of severity and anxiety in the experimental conditions from pre-search to post-search. An ANOVA determined if there were group differences in perceived susceptibility. The fourth section addresses the last two research questions of possible hypochondria and literacy (health and numeric) moderation effects on the relationship of pre-search severity and pre-search anxiety on post-search severity and post-search anxiety, respectively, utilizing two-step hierarchical multiple regressions. The pre-search variable and the possible moderator were entered in Step 1. The interaction effect of the pre-search variable and the moderating variable were added in Step 2. If there was a significant interaction effect, a simple slopes analysis was conducted to see the effect of the relationship.

Prior to analysis, the dependent and demographic variables were simplified and reverse coded questions were recoded. Severity, susceptibility, and anxiety scores were averaged across symptoms to give one score. For the severity and anxiety scores, pre-search (before participants completed the WebMD search task) and post-search ratings were

averaged. In total, eight composite scores were calculated for each participant: severity pre-test, severity post-test, anxiety pre-test, anxiety post-test, susceptibility, hypochondria, health literacy, and numeracy.

Descriptive Statistics

Descriptive statistics of the main variables are presented in Table 3. In each experimental condition, 25 males and 25 females were assigned for a total of 150 participants. The average age of the participants was 19.18 years old and ranged from 18 to 48 years old. Scores from the Whitley Index were used to assess participants' hypochondria with possible scores ranging from low hypochondria starting at 1 and up to 7 indicating high levels of hypochondria. The participants averaged at 2.79. Health literacy was measured by the REALM score. On average, participants did very well with an average accuracy rate of 96%, indicating a high school health literacy level or above. Numeracy, the numerical literacy, was tested using the Berlin Numeracy test and the extended numeracy test by Lipkus et al (2001). Scores represent the accuracy rate of the 12 item numeracy test. Participants correctly answered 75% of the questions.

Severity, susceptibility, and anxiety questions were rated on a scale of 1 to 7. The higher the score, the more participants believed that symptoms were severe, they were susceptible to conditions associated with the symptom, and the more anxiety they had. Severity and anxiety were assessed pre and post web search. Participants usually increased in perceived severity and decreased in perceived anxiety. Susceptibility was assessed by pairing symptoms with conditions found on the web search page. Two conditions were benign and

two were severe. Participants averaged a score of 4.08 on overall susceptibility, with benign susceptibility scores higher than severe susceptibility scores.

Table 2. Descriptive Statistics

	Control		Pictorial		Numeric		All	
	M	SD	M	SD	M	SD	M	SD
N	50		50		50		150	
Gender (F)	25		25		25		75	
Age	18.90	1.43	19.46	3.83	19.18	4.34	19.18	3.43
Whitley Index	2.77	.98	2.89	.90	2.70	.78	2.79	.89
REALM	.97	.07	.97	.03	.98	.03	.97	.05
Numeracy	.74	.13	.74	.15	.73	.15	.74	.14
Working Memory	56.76	13.97	57.46	11.79	61.12	11.50	58.45	12.53
Severity								
Pre-Search	3.79	.99	4.16	1.01	3.99	.88	3.98	.97
Post-Search	3.97	1.04	4.21	.95	4.21	.95	4.13	.98
Anxiety								
Pre-Search	3.82	1.32	3.55	1.34	4.10	1.17	3.82	1.29
Post-Search	3.82	1.39	3.48	1.30	4.03	1.18	3.78	1.31
Susceptibility	3.82	.77	4.04	.87	4.39	.87	4.08	.86
Benign	4.45	.78	4.68	.78	4.98	.94	4.70	.86
Severe	2.92	1.01	3.11	1.20	3.60	1.07	3.21	1.12

Predictor Variables of Post-Search Severity, Anxiety, and Susceptibility

Several variables were found to be predictive of post-search severity, anxiety, and susceptibility. Specifically, the more literate participants were in health and math, the less severe they believed the symptoms were. Additionally, the higher participants rated severity of a symptom before the web search, the higher they rated severity post search. These three variables predicted 79% of the variance in post-search severity.

Similarly, the more literate participants were in health matters the less anxious they reported feeling after the web search. Pre-search anxiety was also highly positively predictive of post-search anxiety. Specifically, the more anxious a participant reported feeling before the web search, the more anxious they felt after completing their search on WebMD. These two variables explained 89% of the variance in post-search anxiety.

Three variables were predictive of susceptibility: health literacy, pre-search severity, and display type. Participants with higher health literacy tended to believe that they were less susceptible to a condition; however, the more severe participants rated a symptom before a web search, the more likely they believed they were susceptible to a condition. Finally, display type was highly predictive of susceptibility with participants feeling more susceptible to conditions if incidence rates were present on a search results page. This relationship will be further explored in the next section. A complete linear regression analysis is available in Table 3.

Display Type Differences on Post-Search Severity, Anxiety, and Susceptibility

There were no differences among the three experimental groups (control, pictorial, numeric) in demographic, literacy, cognitive, and pre-search measures (gender, age, Whitley

Index, REALM scores, numeracy, working memory, pre-search severity, and pre-search anxiety).

Although there were no differences in the severity of symptoms among the experimental groups, participants rated symptoms higher in severity after they completed the web search (*M*=4.13), than before the web search (*M*=3.98, *p*<.001). Participants did not

Table 3. Regression standardized beta values for all post-search assessments

	Severity	Anxiety	Susceptibility		
			All	Benign	Severe
Gender (F)	-.05	-.02	-.07	-.02	-.14
Age	-.01	-.01	-.08	-.07	-.09
Whitley Index	.01	-.05	.01	-.02	.06
REALM	-.09*	-.06*	-.15*	-.11	-.17*
Numeracy	-.10*	.03	-.09	-.10	-.07
Working Memory	.00	-.03	.01	.01	.02
Severity					
Pre-Search	.86***	.01	.40***	.29**	.37***
Anxiety					
Pre-Search	-.06	.95***	.07	.61	.07
IR Presence	-.12	-.04	-.18	-.10	-.25
Display Type	.14	.02	.39**	.32*	.43**
R^2	.79	.89	.30	.19	.29

Note: *p<.05, **p<.01, ***p<.001

differ in anxiety levels from pre-search to post-search nor did they differ among the display groups.

Participants felt most susceptible to conditions when the search results were presented numerically (M= 4.39), followed by pictorially (M=4.04), and the least susceptible to conditions in the control group (M=3.82, F=5.98, p=.003). Susceptibility was further dissected into benign susceptibility and severe susceptibility. Participants felt more susceptible to benign conditions than the severe conditions in all three display groups (F=609.73, p<.001), suggesting they are not escalating after reviewing Internet search results. Furthermore, when examining benign conditions, participants felt more susceptible to conditions when incidence rates were displayed as a percentage in the numeric group (M=4.98) than as a control (M=4.44, p=.002). There were no perceived susceptibility differences between pictorial and numeric or between pictorial and control. Similar trends were seen with severe conditions, with participants feeling more susceptible when incidence rates were displayed as percentages in the numeric group (M=3.60), than the control (M=2.92, p=.003). Unlike the benign conditions, participants also felt more susceptible to conditions when they viewed incidence rates in the numeric group than the pictorial group (M=3.11, p=.028). No difference was found between pictorial and control, suggesting that participants are more likely to escalate when incidence rates are presented as percentages than as bar graphs. Figure 1 below shows the distribution of susceptibility ratings.

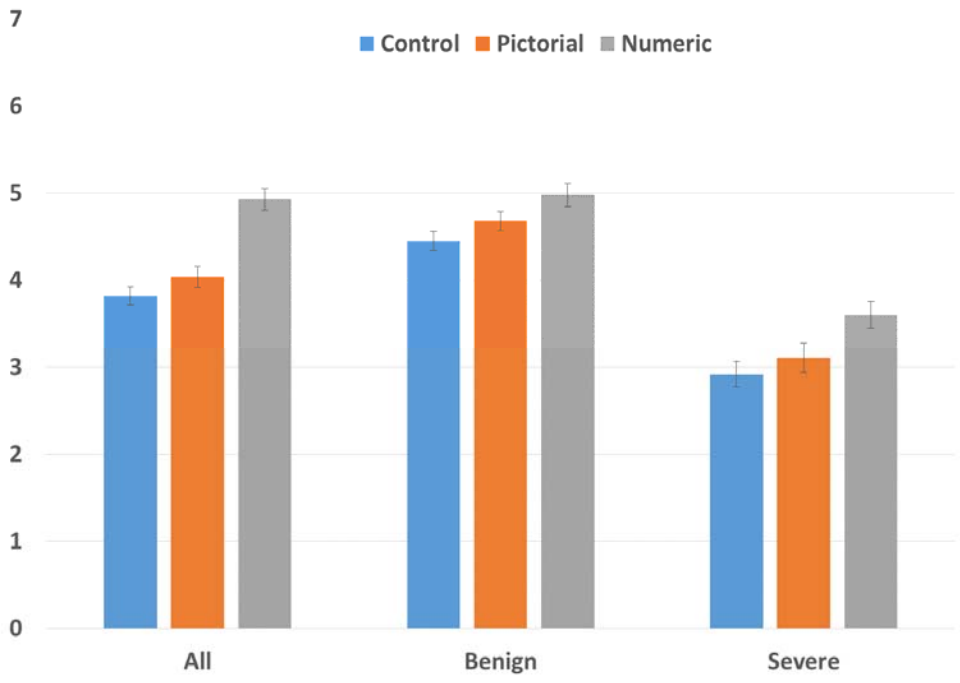

Figure 1. Average ratings for each experimental group for overall susceptibility, benign susceptibility, and severe susceptibility. Error bars represent the standard error of each mean.

These results suggest that the first and second research hypotheses were not supported by our results. There were no group differences between incidence rate presence (present versus not present) or among experimental groups (control versus pictorial versus numeric) groups for severity and anxiety. Susceptibility levels were higher with the presence of incidence rates, specifically highest in the numeric condition, followed by pictorial and lowest in the control, which is opposite to what was hypothesized.

Moderation Effect of Hypochondria and Literacy on Severity and Anxiety

Hypochondria levels, measured by the Whitley Index, had a significant moderating effect on anxiety from pre-search to post-search. The overall trend of the data revealed that participants who rated pre-search anxiety high also rated post-search anxiety high (Figure 2). Similarly, participants who had low anxiety levels prior to the web search also had low

anxiety after viewing the search results page. Upon further analysis of simple slopes, those who had the lowest level of hypochondria and high levels of pre-search anxiety had the highest post-search anxiety as compared to those with mean and high hypochondria levels. Conversely, those with the lowest level of hypochondria and the lowest levels of pre-search anxiety had the lowest levels of post-search anxiety as compared to those with mean and high hypochondria. This phenomenon suggests that Internet searches may lower the anxiety of those who have high hypochondria levels and have high anxiety prior to searching the Internet.

Hypochondria levels did not have a moderating effect on the relationship of severity from pre-search to post-search. Additionally, health literacy and numeracy did not moderate the anxiety and severity relationships from pre-search to post-search. Hierarchical results are displayed below in Table 4. The third hypothesis was partially supported by our results with

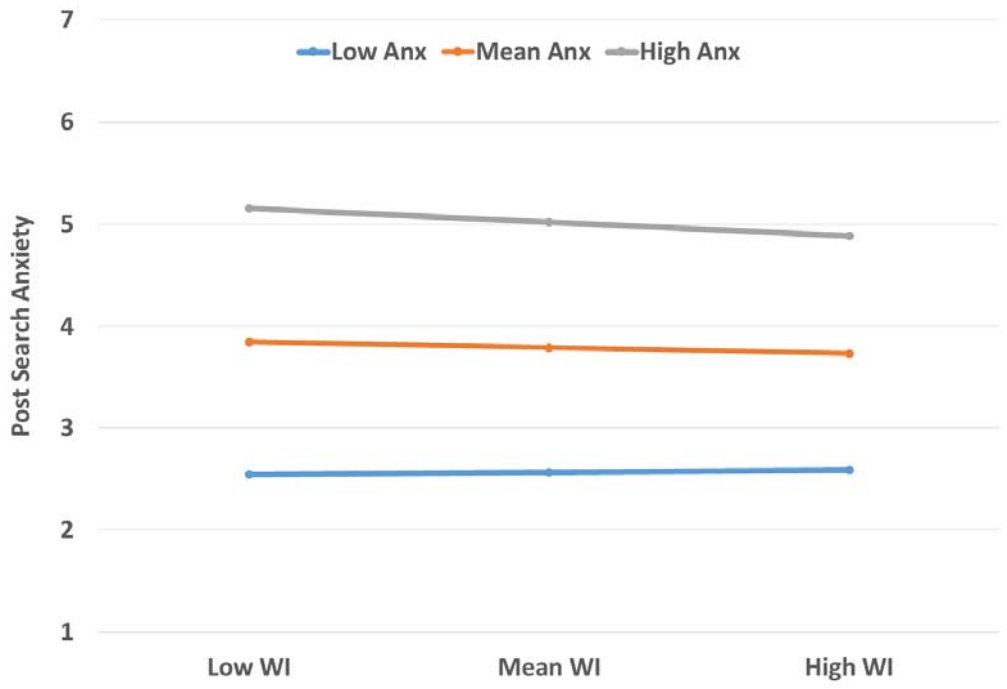

Figure 2. Moderation effect of hypochondria, measured by the Whitley Index (WI), on the relationship of anxiety from pre-search to post-search. Each line represents the pre-search anxiety level.

the moderating effect of hypochondria on the change of anxiety levels from pre-search to post- search. The fourth hypothesis is not supported by our results. Literacy (health and numeric) is not a moderator of the relationship of change in severity and anxiety.

DISCUSSION

Three escalation variables were analyzed in the context of incidence rate presentation for this study: severity, anxiety, and susceptibility. Although severity was shown to significantly increase from pre-search to post-search, susceptibility was the only variable to show group differences in incidence rate presentation. Participants felt more susceptible to

Table 4. Hierarchical regressions examining the moderating effects of hypochondria and literacy (health and numeric) on the change in severity and anxiety.

	Severity				Anxiety			
	β	*p*	R^2	Δ R^2	β	*p*	R^2	Δ *R*
Step 1			.76	.76			.89	.89
Whitley Index	-.00	.935			-.05	.094		
Pre-Search Severity	.87	<.001						
Pre-Search Anxiety					.95	<.001		
Step 2			.76	.00			.89	.00
Whitley Index	-.01	.887			-.04	.126		
Pre-Search Severity	.88	<.001						
Whitley Index*Pre-Search Severity	.03	.443						
Pre-Search Anxiety					.94	<.001		
Whitley Index*Pre-Search Anxiety					-.06	.044		
Step 1			.77	.77			.89	.89
REALM	-.10	.011			-.05	.064		
Pre-Search Severity	.86	<.001						
Pre-Search Anxiety					.94	<.001		
Step 2			.77	.00			.89	.00
REALM	-.11	.027			-.05	.058		
Pre-Search Severity	.86	<.001						
REALM*Pre-Search Severity	.01	.778						
Pre-Search Anxiety					.94	<.001		
REALM*Pre-Search Anxiety					.02	.533		
Step 1			.77	.77			.89	.89
Numeracy	-.11	.010			.02	.580		
Pre-Search Severity	.85	<.001						
Pre-Search Anxiety					.94	<.001		
Step 2			.77	.00			.89	.00
Numeracy	-.09	.026			.02	.463		
Pre-Search Severity	.85	<.001						
Numeracy*Pre-Search Severity	-.04	.316						
Pre-Search Anxiety					.94	<.001		
Numeracy*Pre-Search Anxiety					.03	.297		

conditions when incidence rates were present than when they were not. Specifically, participants felt most susceptible when incidence rates were presented as percentages and the

least susceptible when there were no incidence rates present. Susceptibility ratings were also higher with benign conditions than they were with severe conditions. This indicated that participants believed that the benign conditions were more likely to happen and the severe were less likely to happen. These results also fall in line with the statistical results shown on the search result's page. Previous escalation and cyberchondria studies suggested this phenomenon (Aiken et al., 2012; Starcevic & Berle, 2013; White & Horvitz, 2009), but it has not been tested empirically up until this point.

On the contrary, this study does not follow previous base rate neglect research. Base rate neglect suggests that people are so reliant on heuristics that they neglect base rate or incidence rate data (White and Horovitz, 2009). For example, users could take advantage of the availability heuristic while searching for information online. The availability heuristic occurs when judgments are guided by the ease with which examples come to mind (Tversky & Kahneman, 1973). If a particular disease has been popular in the media, or someone in the family has been diagnosed with a condition, it will be easier to recall than other illnesses. White and Horovitz (2009) suggested this could explain people's failure to take into account the low probabilities of rare conditions in thinking about outcomes and therefore explain why they escalate when searching for health information.

One explanation for the lack of base rate neglect could be an alternate cognitive strategy described by Ajzen (1977). In his study, he found that participants use statistical information only during causal situations. He argued that Kahneman and Tversky (1973) gave base rate information that was non-causal to decision making. Kahneman and Tversky (1973) would give descriptive information about individuals such as "Jack is a 45-year-old man…He is

generally conservative, careful, and ambitious. He shows no interest in political issues and spends most of his free time on his many hobbies which include home carpentry sailing and mathematical puzzles". Participants were then asked to determine if Jack was a lawyer or engineer. They were also given the information that 70% of people in the room are lawyers and 30% are engineers or 30% of the people are lawyers and 70% are engineers. The results of the study found that participants rated Jack at the same rate to be an engineer or lawyer, regardless of the distribution of engineers and lawyers. Researchers concluded that participants were utilizing the representative heuristic and ignoring the base rates provided. The representative heuristic applies when participants are given specific evidence and they choose outcomes based on the representativeness of the evidence. In the study, participants focused on the individuating information of the personality descriptors to make the decision of whether the participant was an engineer or a lawyer, regardless of the base rate. When the extraneous information was taken away, such as in the control condition, participants would utilize the statistical information given (Kahneman & Tversky, 1973).

In the present experiment participants were only given the following description "Imagine that you had chest pain for the past few days and decided to research the information online." There was no extraneous or individuating information such as additional symptoms, family history, what type of pain it was, history of test results such as cholesterol levels, lifestyle descriptions, etc. present. Therefore, participants used the statistical information to make judgments. This is particularly present in the susceptibility question, where participants made judgments on symptoms and the chance of having a particular disease. Examples of susceptibility questions include "If I have abdominal pain, I am at risk

for having food poisoning" for a benign condition and "If I have abdominal pain, I am at risk for having appendicitis" for an escalation condition. Food poisoning was rated at 60% and appendicitis was represented by 20%. This was a causal relationship, therefore, the statistics were used and base rates were not neglected. In fact, participants in the numeric group had the highest susceptibility scores and were significantly higher than the control, which did not have any statistical information.

Hypochondria was a significant moderator for the change in anxiety. Notably, those with high hypochondria levels and high pre-search anxiety had lower post-search anxiety scores than those with lower hypochondria levels. This could suggest that Internet search could lower the anxieties of those with high hypochondria levels. Participants with high anxiety levels could be examining the search results recognizing that benign conditions are more likely than severe ones and, therefore, lowering their anxiety. These results also fall in line with previous research that the presence of incidence rates is lowering anxiety and escalation of participants (Aiken et al., 2012; Starcevic & Berle, 2013; White & Horvitz, 2009).

Neither health nor numeric literacy were moderators for severity or anxiety. This could be due to the fact that the constructs used to measure severity and anxiety were strong and reliable. As seen in Table 5 in Appendix A, the correlation between pre-severity and post severity is .84 and the correlation between pre-anxiety and post-anxiety is .91. This means that 71% and 83% of the variability in severity and anxiety, respectfully, has been explained, leaving a small amount of variance left over to be accounted for by other variables. The moderation effect of the literacies would have needed to be very strong to be significant and this was not present in this study.

Limitations

There are certain limitations that are present in this study. Undergraduate students do not necessarily represent the Internet user population that search for health information online. Undergraduate students are typically young and healthy. They might not have had previous experience with illnesses and therefore, not have much knowledge of health conditions or their consequences. This could also explain why they may not have been anxious evaluating the information.

Additionally, the task presented to the students was artificial. Students were asked to imagine that they were experiencing chest pain, headaches, abdominal pain, and fatigue while conducting an Internet health search. Participants may not have experienced certain symptoms in the past so it may have been difficult for them to imagine experiencing that symptom. Because they were not suffering from the symptoms, they may not have been evaluating the results as seriously as if they actually were experiencing the symptoms. Therefore, they may not have seen the symptoms as very severe or become anxious viewing the search page results and consequently the results of the study may not reflect actual Internet health search behavior.

Additionally, when conducting an Internet health search, you are searching with semantic information, such as how the symptoms began and any previous experience with similar symptoms in the past, whether it be personally related or through someone they know. This information could be distracting in a real world scenario invoking base rate neglect. Further consideration needs to be taken in studying more realistic scenarios to determine if the effect holds.

Future Directions

There are multiple factors that contribute to Internet health escalation. To obtain a more complete analysis of the Internet user population researching health topics on the Internet, there are a few things to consider. First, it would be beneficial to obtain data from a larger, more diverse population. One way could be utilizing a survey website, such as Amazon's Mechanical Turk. This will allow for a larger distribution in age and health experiences, which should provide for more generalizable results.

Additionally, recruiting participants who are actually sick and interested in using the Internet would be valuable. This could be done through a collaboration with the student health center or a local doctor's office. Participants could be asked about their Internet research habits online, and whether or not they used the Internet before coming into the doctor's office, and if they had a preconceived diagnosis before seeing the doctor. Another item that could be analyzed is if the Internet search made them concerned enough to actually come in and visit the doctor. Participants could be asked to perform a search before seeing the doctor and measure their susceptibility, severity, and anxiety for their specific symptoms.

Finally, although perceptual self-report measures were used in this and previous studies to quantify escalation, physiological measures may be able to find effects that participants do not report. For example, participants may report low anxiety while searching for information, but their heart rate and galvanic skin response might be a more demonstrative measure of higher anxiety. Furthermore, eye tracking could be implemented to evaluate if participants are actually looking at incidence rates, and if so how much time is spent examining them.

The Internet is a great source for health information; however, caution needs to be taken on how the information is presented. Escalation can occur especially in the perceived susceptibility of a condition. This study demonstrates that escalation can be minimized by providing statistical information; however, further research is warranted to see if the effect holds during more realistic scenarios and if there are other variables that are in play during escalation.

REFERENCES

Aiken M., Kirwan G., Berry M., & O'Boyle CA. (2012). The Age of Cyberchondria. *Royal College of Surgeons in Ireland Student Medical Journal*, 5, 71-74.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, 35(5), 303-314.

Ancker, J. S., & Kaufman, D. (2007). Rethinking health numeracy: A multidisciplinary literature review. *Journal of the American Medical Informatics Association*, 14(6), 713-721.

Baumgartner, S. E., & Hartmann, T. (2011). The role of health anxiety in online health information search. *Cyberpsychology, Behavior, and Social Networking*, 14(10), 613-618.

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293-332.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: The Berlin numeracy test. *Judgment and Decision Making*, 7(1), 25-47.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175-191.

Fox, S., & Duggan, M. (2013). Health online 2013. Washington, D.C.: Pew Internet & American Life Project. Retrieved from: http://www.pewInternet.org/files/oldmedia//Files/Reports/ PIP_HealthOnline.pdf

Grimes, D. A., & Snively, G. R. (1999). Patients' understanding of medical risks: implications for genetic counseling. *Obstetrics & Gynecology*, 93(6), 910-914.

Hawley, S. T., Zikmund-Fisher, B., Ubel, P., Jancovic, A., Lucas, T., & Fagerlin, A. (2008). The impact of the format of graphical presentation on health-related knowledge and treatment choices. *Patient Education and Counseling,* 73(3), 448-455.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological review*, 80(4), 237-251.

Lauckner, C., & Hsieh, G. (2013). The presentation of health-related search results and its impact on negative emotional outcomes. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 333-342). ACM.

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making*, 21(1), 37-44.

Marteau, T. M., & Bekker, H. (1992). The development of a six-item short-form of the state scale of the Spielberger State—Trait Anxiety Inventory (STAI). *British Journal of Clinical Psychology*, 31(3), 301-306.

Muse, K., McManus, F., Leung, C., Meghreblian, B., & Williams, J. M. G. (2012). Cyberchondriasis: Fact or fiction? A preliminary examination of the relationship between health anxiety and searching for health information on the Internet. *Journal of Anxiety Disorders*, 26(1), 189-196.

Parker, R. M., Baker, D. W., Williams, M. V., & Nurss, J. R. (1995). The test of functional health literacy in adults. *Journal of General Internal Medicine*, 10(10), 537-541.

Raine, L., Horrigan, J., Lenhart, A., Spooner, T., Burke, M., Lewis, O., & Carter, C. (2000). *The online health care revolution: How the Web helps Americans take better care of themselves.* Washington, D.C.: Pew Internet & American Life: Online life report. Retrieved from: http://www.pewInternet.org/files/oldmedia//Files/Reports/2000/PIP_Health_Report.pdf.pdf

Redick, T. S., Broadway, J. M., Meier, M. E., Kuriakose, P. S., Unsworth, N., Kane, M. J., & Engle, R. W. (2012). Measuring working memory capacity with automated complex span tasks. *European Journal of Psychological Assessment*, 28(3), 164-171.

Salkovskis, P. M., Rimes, K. A., Warwick, H. M. C., & Clark, D. M. (2002). The Health Anxiety Inventory: development and validation of scales for the measurement of health anxiety and hypochondriasis. *Psychological Medicine*, 32(5), 843-853.

Seçkin, G. (2010). Cyber patients surfing the medical web: Computer-mediated medical knowledge and perceived benefits. *Computers in Human Behavior*, 26(6), 1694-1700.

Slovic, P., Peters, E., Finucane, M. L., & MacGregor, D. G. (2005). Affect, risk, and decision making. *Health Psychology*, 24(4S), S35.

Starcevic, V., & Berle, D. (2013). Cyberchondria: towards a better understanding of excessive health-related Internet use. *Expert Review of Neurotherapeutics*, 13(2), 205-213.

Top 5 Desktop Browsers from July 2008 to January 2009. StatCounters Global Stats. Retrieved from: *http://gs.statcounter.com/#desktop-browser-ww-monthly-200807-200901*

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, 5(2), 207-232.

Valley, P. (2001, April 28) New disorder, cyberchondria, sweeps the Internet. *The New Zealand Herald.* Retrieved from: http://www.nzherald.co.nz/technology/news/article.cfm?c_id=5 &objectid=185422

Waters, E. A., Weinstein, N. D., Colditz, G. A., & Emmons, K. (2006). Formats for improving risk communication in medical tradeoff decisions. *Journal of Health Communication,* 11(2), 167-182.

White, R. W., & Horvitz, E. (2010). Predicting escalations of medical queries based on web page structure and content. *Proceedings of the 33rd international ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 769-770). ACM.

White, R. W., & Horvitz, E. (2009). Cyberchondria: studies of the escalation of medical concerns in web search. *ACM Transactions on Information Systems (TOIS)*, 27(4), 23.

Wilson, E. A., & Wolf, M. S. (2009). Working memory and the design of health materials: a cognitive factors perspective. *Patient Education and Counseling*, 74(3), 318-322.

Witte, K., Meyer, G., & Martell, D. (2001). *Effective health risk messages: A step-by-step guide.* Thousand Oaks, California: Sage.

APPENDIX

Appendix A

Table 5. Correlation coefficients values (Spearman's rho) between severity (pre and post search), anxiety (pre and post search), hypochondria levels, health literacy, and numeracy.

	Pre-Severity	Post-Severity	Pre-Anxiety	Post-Anxiety	Hypochondria	Health Literacy
Post-Severity	.84***					
Pre-Anxiety	.18*	.13				
Post-Anxiety	.16	.19*	.91***			
Hypochondria	.24**	.19*	.16	.14		
Health Literacy	-.09	-.18*	-.04	-.09	-.15	
Numeracy	-.07	-.13	.01	.05	-.01	.07

Note. *p<.05 **p<.01 ***p<.001