

WHAT'S THE QUALITY OF BREAST CANCER INFORMATION YOU READ ONLINE?

A COMPARATIVE ANALYSIS OF BREAST CANCER INFORMATION QUALITY IN
COMMERCIAL WEBSITES VS. NON PROFIT WEBSITES

A Thesis Presented to the Faculty of the Graduate School

University of Missouri-Columbia

In Partial Fulfillment

Of the Requirements for the Degree

Master of Arts

by

SANTOSH VIJAYKUMAR

Dr. Glen Cameron, Thesis Supervisor

DECEMBER 2005

Thesis Approval Form

The undersigned, appointed by the Dean of the Graduate School, have examined the thesis entitled.

COMPARATIVE ANALYSIS OF BREAST CANCER INFORMATION QUALITY IN
COMMERCIAL AND NONPROFIT WEBSITES

Presented by Santosh Vijaykumar

A candidate for the degree of Master of Arts

And hereby certify that in their opinion it is worthy of acceptance.

Hen T. Cameron

George Devlin

Shelly Boyd

Glen Fisher

Verley P. Pappert

Dedicated To

My mother (Hema)
Father (Vijaykumar)
Sister (Sunitha)
Brother-in-law (Vaidyanath)
Grandaunt (Lalitha)
& Someone who's brought immense
peace and joy to my life over the
past two years, my niece (Nidhi)

ACKNOWLEDGEMENTS

Heartfelt gratitude to my committee members – Glen Cameron, George Demiris, Glenn Leshner, Wesley Pippert and Shelly Rodgers – for all their patience, support and guidance. Each one of you has contributed extensively in making this project an intellectually enriching endeavor for me.

Special thanks are due to George Demiris for help with the reliability testing, Glenn Leshner for guiding the proposal and Shelly Rodgers for her unwavering guidance and encouragement through the analysis phase that was probably the most challenging. Wesley Pippert's experience was invaluable in blending my thesis with the internship at Johns Hopkins University.

My acknowledgement will be incomplete without thanking Glen Cameron, my thesis chairman, for being a fantastic mentor and idea-generator.

Last but not the least many thanks to Ricardo Wray and the faculty at Saint Louis University School of Public Health for their steadfast support through the completion of this project.

Table of Contents

<i>Acknowledgements</i>	<i>ii</i>
<i>List of Tables & Diagrams</i>	<i>v</i>
<i>Abstract</i>	<i>vi</i>
1 Introduction.....	1
2 Theoretical Framework.....	5
3 Literature Review.....	11
3.1 Introduction.....	11
3.2 Overall Scenario of Health Information On the Internet...	13
3.3 Internet Health Information for Varying Ailments.....	17
3.4 Studies comparing the performance of Commercial Vs. Non- Profit Organizations	21
3.5 Literature on Tools for Quality Assessment of Health Related Websites.....	23
3.6 Circumspection.....	26
4 Research Design.....	28
4.1 Explication of Research Problem.....	28
4.2 Description of the Research Methodology.....	31
5 Definitions of Operational Variables.....	35
6 Content Analysis.....	39
6.1 Advantages and Disadvantages of Content Analysis....	42

6.2	List of Research Questions.....	44
7	Results and Findings.....	46
7.1	Factor Analysis.....	48
7.2	Reliabilities.....	50
7.3	T-tests.....	53
7.4	Crosstabulations & Chi-square Tests.....	58
8	Discussions.....	63
8.1	Overview.....	63
8.2	Analyses of the Results.....	64
8.3	Theoretical Implications.....	73
8.4	Practical Implications.....	76
8.5	Recommendations for future Research.....	78
	Appendix 1: Sample Coding Instrument.....	82
	Works Cited.....	85

LIST OF TABLES & DIAGRAMS

Tables

Table 1.....	47
Table 2.....	50
Table 3.....	51
Table 4.....	51
Table 5.....	52
Table 6.....	53
Table 7.....	54
Table 8.....	56
Table 9.....	57
Table 10.....	58
Table 11.....	59
Table 12.....	60
Table 13.....	61
Table 14.....	62

Diagrams

Diagram 1.....	10
----------------	----

A COMPARATIVE ANALYSIS OF INFORMATION QUALITY IN COMMERCIAL AND NON PROFIT BREAST CANCER WEBSITES

Santosh Vijaykumar

Dr. Glen Cameron, Thesis Supervisor

ABSTRACT

Eighty five million Americans access the Internet for health information. But lacks of content regulation, free access, and increased marketing potential have meant that content providers increasingly heed to the call of their own agendas (monetary or non monetary) rather than audience's information needs. An examination of the relationship between the monetary agendas of organizations and the quality of health information on their websites is thus critical in aiding the decision-making skills of online health information seekers.

Considering that breast cancer is the leading cancer site among American women, we conducted a comparative analysis of online breast cancer information between commercial and nonprofit websites. Content in top 20 popular breast cancer websites (comprising a total of 627 web pages) was analyzed using the European Union's quality criterion for online health information. Nonprofit websites had a significantly higher overall score and stated sources with source credentials more frequently than commercial websites did. Commercial websites scored marginally higher in stating funding sources and offering 'current' information. Neither of the two groups made their website handicapped-accessible.

These results highlight the worryingly low quality levels of online breast cancer information, and make a call for increased scholarly and industry attention to this area. This study makes recommendations for future research with due consideration to the most recent explosion of health blogs, and Information Communication Technologies (ICT's) in developing countries in the context of online health information seeking behavior.

CHAPTER 1: INTRODUCTION

The Internet has become one of the most powerful and widely used means to communicate and disseminate information over the past five years, at least. And a major ramification of this is that websites have now become an inextricable tool for both information givers who can use this as an effective mouthpiece and audiences who can download a wealth of information with a mouse-click.

Health is undoubtedly one of the many fields to have seen a rapid proliferation of websites. The information carried on these sites spans the entire spectrum of issues ranging from online health education, to diet consultancies, telemedicine intervention arrangements, health campaigns, and even e-treatments. The Internet is rapidly bridging the gap between information that's designed to be used by health professionals and information that's designed to be used by consumers - Many consumers are eager to have access to all information relating to their health situation, and many of them are quite capable of understanding it (Rumsey).

Health sites are among the most popular searches on the web, with more than 85 million people using the Internet for medical information in the US alone (Losken, et al., 2005). In the backdrop of this health information-seeking scenario it is important to see which diseases have figured most prominently as being searched for by users. A study by Ziebland, et al. showed that breast cancer is among the most searched site of cancer by women.

But try keying in “Breast Cancer” as a search term on Google and chances are that you will be offered at least 77,000,000 sites to choose from for obtaining the information you need. The net has now evolved into as much of a forum for free expression as it is a vehicle of information dissemination. And this means that the health information carried on e-space is not just that carried by federal health agencies but also by corporate bodies that form part of the health industry caucus and various individuals – ranging from professionally qualified medical experts to quacks looking to make a fast buck. Everyone (who wishes to) can have a website giving information on a particular disease. But whose information can the user trust?

Whose information is well researched, well attributed, and credible? This is where a comparative evaluation of websites carrying health information on the Internet becomes critical. For, we are dealing with not just plain information, but information that might well decide lives and certainly the quality of life. Also, it is known that sites with commercial interests have more means of attracting visitors, leaving non-profit organizations with even less chance of reaching users (Hargittai, 2000).

This poses the larger question whether the monetary agendas of health organizations reflect on the health-related content that is carried on their websites. An initial attempt to answer one facet of this question is by studying, analyzing and comparing the quality of websites of organizations that fall into two prominent categories of monetary agendas- commercial and not-for-profit.

This study thereby aims to conduct a comparative evaluation of the overall quality of commercial websites versus those of not-for-profit websites that carry cancer-related information.

In line with this, the research study attempts at answering two over-arching research questions:

1. Which of the two groups of websites – commercial and nonprofit – have better quality of online breast cancer information?
2. What is the nature and extent of the difference between each of the quality parameters when applied to both the website groups individually?

These are the two main overarching research questions. Given the theoretical construct used for this study, we also attempt to examine as to:

3. Which of two groups of websites have a stronger presence of information related to individual constructs from the Health Belief Model?
4. Which of the two groups of websites has a stronger presence of: a) any kind of breast cancer information, and b) commercial information?
5. Among websites that have any kind of breast information, how many of their web pages provide information on the source and credentials of the source of the respective information?

6. Is there a difference between the two groups of websites in being certified by the Health On Net Foundation code of conducting and carrying their seal? If so, what is the extent of this difference?

The 'quality' mentioned above would be evaluated by carrying out a detailed content analysis wherein different aspects of these websites will be pitted against the quality standards for online health information that have been set by the European Union (EU) with consultation from the Health On The Net (HON) Foundation and other organizations (www.hon.ch).

Apart from the apparent addition of knowledge to the realm of web-based medicine and health informatics, I believe that my thesis has the ability to bring forth certain recommendations for quality enhancement for these websites. Also it has the potential to set the tone for future research that specifically studies the agenda-setting and gate-keeping phenomena in health organizations with different monetary objectives that use the Internet as a predominant business component.

CHAPTER 2: THEORETICAL FRAMEWORK

The research statement and objective stated above spreads its wings across different disciplines and theoretical fields like health and medicine, mass communication, new media and nonprofit businesses. Two approaches for integrating research with theory were considered while drafting the theoretical framework.

First, an inside out approach was considered where in we pick aspects from theoretical approaches in each of these different fields and chart out one framework. The second is the outside-in approach where we can choose a theory in the field of Health Communication and fit different aspects of our research question- like the cancer aspect, the new media aspect, the human (patient) aspect and finally the health behavior aspect- and integrate them into one theory. Though the first approach would have resulted in varied perspectives from varied scientific realms, it was decided that the second approach would aid in more cohesion and focus for the purpose of this research project.

A number of theoretical approaches and models serving this purpose was considered: the Extended Parallel Process Model (EPPM), the Protection Motivation Theory, the Theory of Reasoned Action, the Elaboration Likelihood Model and the Social Cognitive Theory. However, it has been decided to adopt the Health Belief Model (Janz & Becker, 1984; Rosenstock, 1974) and integrate the aspects of our research question with the same.

One of the most widely used models of health behavior change (Witte et al, 2001) the Health Belief Model (HBM) suggests that preventive health behavior is influenced by five factors:

- a. Perceived barriers to performing the recommended response;
- b. Perceived benefits of performing the recommended response;
- c. Perceived susceptibility to a health threat;
- d. Perceived severity of a health threat;
- e. Cues to action.

The most interesting considerations in the selection of the Health Belief Model are that its constructs are structured in a way that our study can be operationalized in either of these two manners:

1. The cumulative effect of all the five constructs eventually leading to a behavior – that is accessing online health information, OR,
2. Online information seeking as a route to obtaining detailed information about susceptibility, severity and cues to action, not to mention treatment, which will be evaluated as a separate construct later on.

We can explain the first method mentioned above by positing that the threat in this case being cancer, one of the recommended responses to this threat is to access cancer-related websites and gain information from them.

The HBM suggests that individuals weigh the potential benefits of the recommended response against the psychological, physical, and financial costs of the action (Janz & Becker, 1984; Rosenstock, 1974). Led by a reasonable assumption that today's consumers are discernible and cognizant about their rights and choices, we consider the overall quality of websites as being one of the barriers for consumers not accessing cancer-related websites. This means, given the fact that corporate or nonprofit websites do not meet each one or all of the parameters for quality standards that have been set by the HON Foundation; this will act as a deterrent to health consumers and cancer patients and their relatives from hitting these websites.

Further, the primary stakeholder (the cancer patients) will weigh the benefits of accessing these sites against their barriers and then make a final decision. The benefits could well range from getting tips for self-care to customized daily routines for diet and exercise and other treatment solutions. If these benefits outweigh the barriers, the patient will access the site and make most of the utility being offered. In case of the barriers outweighing the benefits, the entire utility of the content on these websites will be negated by a lack of quality. The quality parameters, by the way, are: credibility, content, disclosure, links, design, interactivity and caveats.

What also needs to be mentioned is the fact that the measurement of benefits by the patient is also a function of her susceptibility and severity towards cancer. For example, the benefits of certain websites might be extremely crucial

and high to a patient whose cancer is at an advanced stage, while it might be of little utility to someone with no advanced cancer.

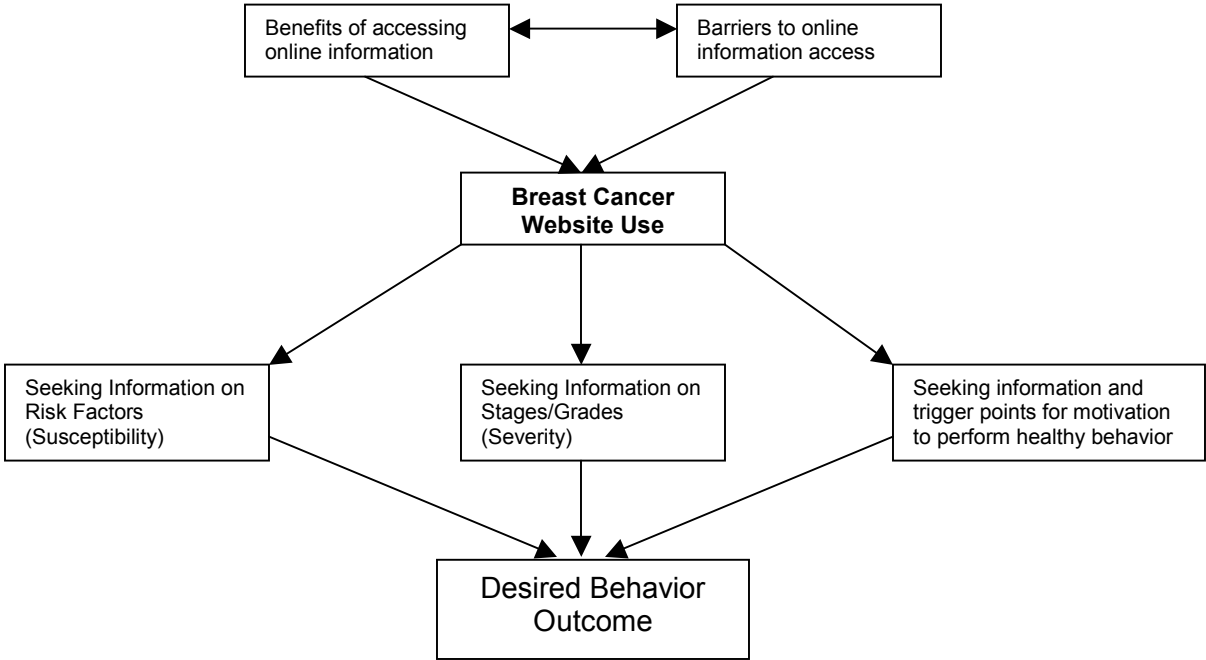
As has been mentioned above, the recommended health behavior change action is also a function of the patient's susceptibility and severity towards the larger threat: cancer. The susceptibility factors to cancer include: family history, congenital abnormalities, pre-cancerous conditions, predisposing diseases, exposure to carcinogens like smoke and tobacco that induce cancer and viral and paralytic infections. Evaluating either the 'stage' or the 'grade' of cancer, however, can assess the severity of cancer. While the stage is determined by how far and wide in the body the cancer has spread, the 'grade' of cancer is a function of loss of body cell differentiation.

Finally, the "cues to action" are defined as internal or external factors that trigger the subject into performing the recommended health behavior by increasing the perceptions of susceptibility and severity. For example, a patient who reads a certain cancer magazine might be led to accessing a cancer-related website in one of the articles in the magazine to obtain more in-depth cancer information.

Thus, this method warrants the use of more resources – in terms of personnel, finances and logistics – in order to study how each of the five constructs translates into the final behavior of obtaining online information.

For purposes of this study, we have this adopted the second method where we will choose three of the five constructs from the model – Susceptibility, Severity and Cues to action – and, by using these as quality parameters, evaluate the performance of the two groups of websites as to whether they address these issues on their sites by providing relevant information about each. The three constructs will thus be used as quality parameters and a quantitative evaluation and analysis carried out correspondingly. A diagrammatic representation of this theoretical framework is illustrated below.

Diagram 1: Theoretical Framework



CHAPTER 3: LITERATURE REVIEW

3.1: INTRODUCTION

A comparative evaluation of the overall quality of commercial versus nonprofit cancer websites cuts across realms of health and medicine, mass communication and modern business. It is common knowledge that each one of these realms is undergoing rapid developments, both in terms of technological advancement as well as the nature of academic research being conducted. A research study conducted towards the evaluation of cancer websites will thus be incomplete without a detailed review of the scholarly work that has been performed in different facets of this examination area. It is also important to note that the literature under review must be carefully chosen to be the latest in this field, lest the review remains outdated.

Considering that there are many different facets to this research study, we have decided to compartmentalize the chosen literature into the following categories and later, review their significance and importance in the context of our project:

- Overall scenario of health information on the Internet; its procurement; its dynamics.
- Internet information for ailments other than Cancer, like Alzheimer's disease, Depression and Arthritis.
- Examinations of quality indicators for health information on the Internet.

- Impact of business agendas on the performance of organizations and their websites – as in, how being a nonprofit or commercial entity impacts on the organization's website.

3.2: OVERALL SCENARIO OF HEALTH INFORMATION ON THE INTERNET

A vast amount of scholarly work, it has been found, has been devoted to studying the role of Internet in the world of healthcare and medicine. Though there have been studies that have examined this role solely in the context of the health consumer, there has also been a sufficient quantum of work employed towards detailing this in the context of industry, physicians and other healthcare professionals. It is interesting to see how this research has been quickly complemented by further work in delineating the most potent tools for procuring quality information on the Internet. Needless to say, some of these tools also direct the user to websites disseminating information of spurious quality. This has given rise to greater concerns, the most paramount of which is to draft a mechanism by which the quality of health related information can be assessed, and standards be set.

Observing that the Internet has the ability to greatly affect and improve the quality of healthcare, Kerwin & Madison (2002) also note that the growing interest toward Internet related healthcare activities is causing significant changes in the U.S. healthcare environment. “Although the movement of the healthcare industry to Web-enabled systems and processes has been somewhat slow in comparison to the movement in other industries, explosive growth in the rate of adoption of Internet technologies by the healthcare industry is taking place.”

Kerwin & Madison also examined the economic viability of the Internet and posited that it has proved to be a prominent cost-cutting instrument in the scenario of the health business. No wonder then, that e-health is now proliferating as rapidly as each of its components - new media and health. An integral and interesting part of Kerwin's study also details on how the concept of Internet has pervaded into organizations in ways like intranets and extranets resulting in more efficient communication between healthcare employees. Furthermore, the study also touches on the use of e-mail for physician-patient communication and discusses the position of patients as "healthcare consumers".

In what seems like having taken one strand from Kerwin's study for detailed exploration, Domaszewicz (2002) delineates the online tools used in the paradigm of human resources for healthcare. Domaszewicz's position is that the Internet has enabled the creation of a new generation of tools to help employees make healthcare decisions based on more than per-paycheck deductions and word-of-mouth quality. This assumes more significance when pitted against the larger backdrop of patient awareness of the quality of healthcare they are provided.

Putting this into perspective, Domaszewicz moves on to detail the use of online tools like plan selection tools, and those used for comparing and eventually choosing a healthcare provider. These generic tools of course have other more specific tools embedded in their systems, thereby giving the final consumer the gift of choice and eventually, a well-informed decision.

Interestingly, prior research has been conducted on a new generation of tools for search, recovery and quality evaluation of World Wide Web resources. Noting that the 'portal' is becoming a one-stop shot for information resources on the Internet, Aguillo (2000) prepared a new portal on health-related resources mainly devoted to Spanish-speaking countries. It is during this preparation that Aguillo faced a number of challenges, which he delineated in his study. According to Aguillo foremost of this set of conundrums was that of information overload on health-related Internet. Observing that the size of the web increases exponentially (over 250% per year), he noted that the current Internet search engines to scan its contents would be a serious handicap. This is an extremely important observation in the context of our project for two reasons: one is that, the author of this research project proposed the use of search engines in order to procure all of the corporate and nonprofit websites for the purpose of this project; and the second being that this also puts a quick thought on the sifting through and selection of health websites by patients and other health consumers.

Amidst a sea of other indicators and parameters that Aguillo used for this project, he concluded that the critical nature of medical information results in the requirement of new techniques and methods in order to guarantee a portal's viability, adequate coverage of new resources and quality of its contents. Finally discussing the area of quality evaluation, Aguillo suggests, "The quality evaluation involves the building of a first or original sample of super-sites, selected and indexed by human experts and covering a wide range of medical and health-related topics."

However, the wealth of health information on the Internet does not come without its own share of snafus and snags. The advent of spam, it has been identified that there has been a growing number of “dubious” health claims on the web pages (Cooke). These include advertisements with claims that dehydroepiandrosterone (DHEA) helps prevent or improve cancer, obesity, Alzheimer’s ageing, cholesterol and brain function, and that Melatonin “*strengthens the body’s immune systems*”. While the US FDA does not approve DHEA, Melatonin has been banned in UK. Cooke’s study underlined similar chronic problems associated with Internet health information and details problems with gateways to high quality Internet-based resources. Marking the need for the quality of internet health info, Cooke concludes, “The importance of evaluating the quality of information is nowhere more apposite than in relation to the Internet - anyone anywhere can publish with the simple proviso that they have access to the relevant hardware and software and, in principle, anyone can pretend to be whoever they want to be.”

It is important to note in the context of our study that Internet health information has been an active topic of discussion among health and mass communication scholars. The content of these studies clearly direct us to the fact that though the Internet is almost basic to healthcare delivery in today’s context, there is a fervent need to monitor the content that is carried on the expanding domains of the Internet.

3.3: INTERNET HEALTH INFORMATION FOR VARYING AILMENTS

An apparent extension to a review of the literature that suggests how swamped health consumers are with Internet information is by analyzing studies that have been conducted on the use of online info across various ailment groups. The literature review conducted on this subsection of the paper suggests that though there have been independent studies conducted on quality of online information for Alzheimer's disease, depression and specific kinds of Cancer, there has been very little work conducted on the larger domain of cancer-related websites. This is the hole in academic research that our research project attempts to fill. However, this does not in any way discount the efforts and effects of studies conducted on these other diseases. Because, it is through the prism of these studies that we find a circumspective angle to conducting quality evaluation studies on areas that we intend to, in this project. Alternatively, these studies also direct us towards considering ramifications of research in this area and guides us towards pursuing different, yet detailed work in the future, like carrying out a comparison of the quality of Cancer websites versus those of other ailments.

In his study conducted on the quality of web-based information on treatment of depression, Griffiths (2000) conducted a cross-sectional survey to discover that the overall quality of depression websites was poor- a paltry 4.7 on a range of 0-13.

The study, which was aimed to survey websites that a ‘typical’ user might access when searching for information on depression, used two search engines, DirectHit (www.directhit.com) and MetaCrawler (www.go2net.com/search.html). It is important to note that while DirectHit returns 10 “popular” sites based on analyses of previous user activity for a query, MetaCrawler integrates the results for a query from at least nine other well-known search engines including Yahoo and AltaVista. Furthermore, the quality analyses were conducted based on three parameters: characteristics, content and accountability. Interestingly, Bath used the evidence-based guidelines on clinical practice for treating depression published by the Agency for Health Care Policy and Research (ACHPR).

The study concluded that, “There is a need to improve the accuracy and coverage of information about depression on the web with regard to the relative effectiveness of different treatments, the main indications for particular treatments, important management issues such as duration of treatment, reviewing and changing treatments, and the relevance of professional expertise and patient preferences”.

The quality of Alzheimer’s websites too did not provide an extremely encouraging picture. In a study conducted in this direction, Bath (2000) compared the quality of websites assessed by the Alzheimer’s disease evaluation tool against that assessed by a new evaluation tool that was created especially for this study.

Using a scoring system, the tool made use of parameters like general details, currency, Carers’ Information, Using the Website and Conclusions in

order to carry out their examination. Apart from some interesting results that this study has given, the most important conclusion derived was that websites generally scored lower using the specific tool compared with one of the generic tools. Bath opined that “further work is required to develop and evaluate the Alzheimer’s disease website evaluation tool to improve the quality of information available to Carers’ through the Internet.

Zielbland et al. (2004) conducted one of the more relevant studies on online health information use that examined the effects of Internet on patients experience with cancer. Hundred and seventy-five cancer survivors, who had been diagnosed in the past 10 years, were studied in order to find out presence and patterns of Internet use to seek online information, and also perceptions about cancer after obtaining information online.

It was found that most women sought information on breast cancer with a specific emphasis on breast cancer treatment. Additionally, it was also found that the kinds of cancer information sought online depended upon what stage of diagnosis or treatment the patients were at. For example, patients would use the Internet to find out possible meanings of their symptoms before seeing the doctor the very first time; contrastingly, they would seek information about alternative treatments, research, and experimental treatments in the first stages after diagnosis.

Finally, and most significantly in the context of this research project, Meric et al (2002) conducted one of the few scholarly studies on quality of breast cancer information on the Internet. The study not only examined if websites met

certain quality criteria but also the attributes of the content. Within the total sample of 185 websites, the authors identified 84 commercial and 64 nonprofit websites. Various attributes of quality – display of authorship, attribution, currency, disclosure, HON code seal and e-mail address were studied to compare more popular and less popular websites. Commercial sites were found to contain more inaccurate medical statements than nonprofit websites and websites of other organizations. Interestingly, though more commercial websites displayed the HON Code seal, three of them commercial websites contained inaccurate medical statements.

3.4: STUDIES COMPARING THE PERFORMANCE OF COMMERCIAL VERSUS NONPROFIT ORGANIZATIONS AND THEIR WEBSITES

A logical continuation of our previous section will be briefly reviewing the scholarly work that has been conducted on the different dynamics of the performance of commercial and nonprofit organizations and their websites. Thorough and repeated searching attempts to excavate studies that have conducted a comparative evaluation of commercial and nonprofit health websites gave only one result – this was a study conducted by Kunst and Khan (2002) comparing the quality of medical information on COPD provided by commercial versus non-commercial websites. The study revealed some interesting results like the presence of substantial gaps in the credibility and accuracy of the content of websites on stable COPD. These gaps, says Kunst, were significantly greater among commercial sites compared to non-commercial sites.

Apart from that however, there is little consolation. This study seems to be a lone ranger in this area that offers so much scope for exploration, discovery and analysis in the vast ocean of e-health.

However, there have been certain studies whose insights might be peripheral for our study, but they are extremely useful nonetheless. One such study is a huge research initiative undertaken by Rosenau (2003) where he conducted a performance evaluation of for-profit and nonprofit U.S Hospitals since 1980 based on data collected between 1980 and 2002. The study said that nonprofit hospitals performed better than for-profit hospitals than the contrary. It

will be intriguing to hypothesize whether the ramifications of this result can be extended to the performance of the websites of the respective groups. Also, Rosenau brings to light some “light theoretical controversies”. The most important of these in the context of our study being that the debate about which provider type is better, for-profit or nonprofit, may never be resolved to everyone’s satisfaction.

There has also been some work carried out in comparing the two groups in a totally different field – real estate. In a content analysis comparing commercial real estate brokerage and residential real estate brokerage websites, Henderson and Cowart (2002) examine the structure of these websites with the goal of determining whether some patterns of content might increase the quality and quantity of information available to buyers and sellers thereby contributing to the disparity between growths in the two groups.

This study offers interesting results in that the authors reveal that residential real estate websites provide more informational content than commercial real estate websites.

3.5: LITERATURE ON TOOLS FOR QUALITY ASSESSMENT OF HEALTH-RELATED WEBSITES

The review of literature spanning the above three dimensions of this research project must suitably culminate with a detailed view of the research that has been conducted on the different instruments and methodologies which modern day researchers use in order to evaluate websites. This is possibly the most crucial part of the this literature review since the crux of this research project lies in evaluating the quality of two different groups of websites based on a quality indicator that, as we have seen, as been often considered and reviewed in previous research work. Again, it is encouraging to see that efforts have been taken in the direction of evaluating the quality indicators that have been set up so far, and what is more encouraging, is that professional organizations of health informatics personnel, like HIMSS, are working towards developing new quality standards for Internet health information.

In one of the most detailed studies conducted in this area, Gagliardi and Jadad (2002) carry out an examination of instruments used to rate Internet health information. Updating work on a previous study conducted in 1998, Gagliardi selected 98 instruments used to assess the quality of websites in the past five years and found that of 51 newly identified rating instruments only five provided some information by which they could be evaluated.

“Many incompletely developed rating instruments continue to appear on websites providing health information, even when the organizations that gave

rise to them no longer exist. Surprisingly, many of these rating instruments of questionable utility and without association to operable entity are featured on the U S Department of Health and Human Services Healthfinder website...which uses a detailed and rigorous selection process for the development of its own content.”

The researchers also discovered that a large number of researchers, organizations and website developers are exploring alternative ways to help people find and use high quality information available on the Internet.

Not surprisingly, the quality of health information too does not seem to provide a reason to celebrate. As Gretchen et al (2001) discovered in their assessment of accessibility, quality and readability of health info on the net in English and Spanish, 24% of the clinical elements on English- and 53% on Spanish language websites were not covered at all. Also, all English and 86% of Spanish websites required high school level or greater reading ability. This study too provides a discouraging conclusion saying that the *“coverage of key information on English and Spanish language websites is poor and inconsistent although the accuracy of the information provided is generally good.”*

However, we have only reviewed the quality assessment of health websites from the designers and/or health professional’s point of view. Healthcare consumers and patients are the most critical stakeholders in the healthcare system and it is crucial to examine what they perceive to be the top quality indicators and discover the bases on which they assess Internet health information. Eysenbach and Kohler (2002) used focus groups, usability tests and

in-depth interviewing methods to find out how consumers search for and appraise health information on the Internet.

The study was conducted on 17 participants who were given a series of health questions and observed in a usability laboratory while retrieving health information from the web. The research ended up offering some fascinating results – though the search technique of these users was often sub-optimal they successfully found health information to answer questions in an average of five minutes and 42 seconds. More importantly, the focus group participants said that their main assessment of credibility of websites was based on whether they found a source, a professional design, a scientific or official touch, language and ease of use. The participants hardly checked the “About Us” link and also, very few participants remembered the exact websites from where they had retrieved their link.

3.6: CIRCUMSPECTION

It can be safely said that this literature review has certainly been successful in empowering the course of this research project with some extremely interesting additional dimensions and multi-perspective insights.

Though there has been work conducted on the above four facets of this research project independently, there are two crucial points to be considered. One, is that the amount of scholarly work conducted on e-health is either too niche or too broad, but it is insufficient either way. Secondly, it is encouraging to perceive that this research project appears to successfully amalgamate the above four facets into one study that will end up offering new insights, add more dimensions and contribute to the overall body of knowledge in this area. If implemented with due caution of bias and error, this project has the potential to fill up some gaping holes that will only result in making the research path a lot smoother.

It is very encouraging to note that established researchers before have already used the HON Foundation's code of conduct. And what is equally encouraging is the fact that the two groups – Commercial and Nonprofit- under consideration for this study offer interesting disparities in quality levels, thus making this study even more necessary.

It is important to conclude from the literature review that this study is by far the first study that conducts a detailed examination and comparative analysis of the quality of the ten most popular commercial and nonprofit breast cancer websites.

The endeavor is to not only add a new dimension to the present body of knowledge on e-health, but also derive perspectives that will empower future researchers to conduct more varied, detailed explorations in this field.

CHAPTER 4: RESEARCH DESIGN

4.1: EXPLICATION OF RESEARCH PROBLEM

The principal aim of this research project is to carry out a comparative evaluation of the overall quality of commercial websites versus not-for-profit websites that carry cancer-related content.

It is known that sites with commercial interests have more means of attracting visitors, leaving non-profit organizations with even less chance of reaching users (Hargittai, 2000). It can thus be deduced that a relationship does exist between the monetary agendas and situations of health organizations and the performance of their websites. This research is a further extension of this deduction. By carrying out a detailed quantitative content analysis of the various attributes of cancer-related websites owned by corporate companies and not-for-profits, we can indirectly deduce if the monetary agendas of these organizations have a bearing on the overall content and quality of the websites.

But how can we measure a website's quality? In order to carry out this evaluation, we are using the quality parameters for online health information established by the European Union and adopted by the Health on Net foundation. These parameters outlined are best explained in this boxed excerpt from the website: http://www.hon.ch/HONcode/HON_CCE_en.htm

“Quality Criteria for Health Related Websites

Developed in widespread consultation with representatives of private and public e-health websites and information providers, other industrial representatives, public officials, and representatives of government departments, international organisations, and non-governmental organizations:

“These criteria should be applied in addition to relevant Community Law

Transparency and Honesty

- Transparency of provider of site – including name, physical address and electronic address of the person or organization responsible for the site (see Article 5 and 6 Directive 2000/ 31/ EC on Electronic Commerce).
- Transparency of purpose and objective of the site
- Target audience clearly defined (further detail on purpose, multiple audience could be defined at different levels).
- Transparency of all sources of funding for site (grants, sponsors, advertisers, non-profit, voluntary assistance).

Authority

- Clear statement of sources for all information provided and date of publication of source.
- Name and credentials of all human/ institutional providers of information put up on the site, including dates at which credentials were received.

Privacy and data protection

- Privacy and data protection policy and system for the processing of personal data, including processing invisible to users, to be clearly defined in accordance with community Data Protection legislation (Directives 95/ 46/ EC and 2002/ 58/ EC).

Updating of information

- Clear and regular updating of the site, with date of up-date clearly displayed for each page and/ or item as relevant. Regular checking of relevance of information.

Accountability

- Accountability – user feedback, and appropriate oversight responsibility (such as a named quality compliance officer for each site).
- Responsible partnering – all efforts should be made to ensure that partnering or linking to other websites is undertaken only with trustworthy individuals and organisations who themselves comply with relevant codes of good practice.
- Editorial policy – clear statement describing what procedure was used for selection of content.

Accessibility

Accessibility – attention to guidelines on physical accessibility as well as general findability, searchability, readability, usability, etc.”

4.2: DESCRIPTION OF THE RESEARCH METHODOLOGY

Given that the area of online breast cancer information quality research is still in the developing stages, the challenge was to come up with a methodology that would adequately address the following issues:

- a. Offer an adequate sample size, whose findings would have a reasonable level of external validity at least within the niche of breast cancer websites
- b. Quality parameters that were easy quantifiable
- c. Avoid overlap among variables (quality parameters) that were being tested, and,
- d. Lend them into offering a reasonably accurate picture of overall quality as outlined by the EU.

Given these challenges, this was the final methodology used for the study:

1. **Determining sample size:** With due to consideration logistical and monetary limitations, we decided to conduct the study on ten websites from each of the two categories – commercial and nonprofit. Upon selection of these twenty sites as the units of assignment, we decided to download all the first level web pages (all links visible on the home page). 'Web page' would thus be the final unit of analysis.

2. **Selecting the sample:** After determining the sample size, we decided to use the purposive sampling technique by selecting the top ten websites from each category. In order to get to this stage, we used Google's 'Advanced Search' facility in order to define the domain (.com for commercial websites and .org for nonprofit websites) and employed 'Breast Cancer' as the search string for getting the ten most popular websites for each category. Google uses their Page Rank algorithm to list their search results on the basis of 'link popularity'. A description of Google's search mechanism that offers the most popular search results upfront is given below (Google, 2005).

“Technology Overview

The software behind Google's search technology conducts a series of simultaneous calculations requiring only a fraction of a second. Traditional search engines rely heavily on how often a word appears on a web page. Google uses PageRank™ to examine the entire link structure of the web and determine which pages are most important. It then conducts hypertext-matching analysis to determine which pages are relevant to the specific search being conducted. By combining overall importance and query-specific relevance, Google is able to put the most relevant and reliable results first.

- **PageRank Technology:** Page Rank performs an objective measurement of the importance of web pages by solving an equation of more than 500 million variables and 2 billion terms. Instead of counting direct links, PageRank interprets a link from Page A to Page B as a vote for Page B by Page A. PageRank then assesses a page's importance by the number of votes it receives. Page Rank also considers the importance of each page that casts a vote, as votes from some pages are considered to have greater value, thus giving the linked page greater value. Important pages receive a

higher PageRank and appear at the top of the search results. Google's technology uses the collective intelligence of the web to determine a page's importance. There is no human involvement or manipulation of results, which is why users have come to trust Google as a source of objective information untainted by paid placement.

- **Hypertext-Matching Analysis:** Google's search engine also analyzes page content. However, instead of simply scanning for page-based text (which can be manipulated by site publishers through meta-tags), Google's technology analyzes the full content of a page and factors in fonts, subdivisions and the precise location of each word. Google also analyzes the content of neighboring web pages to ensure the results returned are the most relevant to a user's query.”

3. **Creation of the coding protocol:** Each of the indicators of separate constructs under the European Union's quality parameters were considered as individual binomial variables. Following this, three of the five constructs as elucidated by the Health Belief Model – susceptibility, severity and cues to action – were added to this list. In line with our focus on the monetary agendas of websites, we decided to add two more variables to the list of measures – information and commerciality. Treatment and HON Code were later added in order to study if websites provided breast cancer information and also, if they carried the HON Code seal on their websites. The operational definitions of each of these variables have been given in the next chapter. This coding protocol was created as a form on MS- Access software and interlinked to tables so that they could later be exported to MS-Excel as a complete dataset. Each of the websites was identified with an alphabet starting A to T. Thus,

websites A to J comprised 10 commercial websites, and K to T where nonprofit websites. Each web page on the homepage of the respective websites was downloaded and saved as an .html file along with graphics. Advertisements were saved too, since they would reflect on the amount of commercial content.

Inter-rater Reliability: Inter-rater reliability, which addresses the consistency of the implementation of a rating system, was explored by inviting a second rater to code a small subset of the web pages (3% of total sample, N=20). Percent of agreement between the two raters was very high (93%) across all variables, which indicates satisfactory levels of inter-rater reliability for the coding protocol. The coding of all the 627 web pages was carried out in a single week to maintain greater consistency.

CHAPTER 5: OPERATIONAL DEFINITIONS OF TERMS & VARIABLES

1. **Quality:** The degree or grade of excellence of a particular website.
2. **Corporate organization:** Body politic, a corporation or a body of persons who carry out a business in order to achieve a certain monetary goal/ monetary profit.
3. **Nonprofit Organization:** A body comprising a group of like-minded people carrying out philanthropic activity for a social cause with no monetary objectives (for the purposes of this study, nonprofit websites were chosen on the basis of whether they were covered under the 501(c) (3) act.
4. **Website:** Defined as a document or a set of linked documents, usually associated with a person, organization or topic that is held on a computer system and can be accessed through the World Wide Web.
5. **Cancer:** Defined as a general term for more than 100 diseases that are characterized by an uncontrolled, abnormal growth of cells.

DEFINITIONS OF VARIABLES

1. **Name and Address:** Name, physical address and electronic address of the person or organization responsible for the site
2. **Purpose:** Purpose and objective of the site
3. **Target Audience:** Target audience clearly defined
4. **Transparency:** Transparency of all sources of funding for site (grants, sponsors, advertisers, non-profit, voluntary assistance)
5. **Sources:** Clear statement of sources for all information provided and date of publication of source (even a single source provided for overall information on the page should be coded as 'yes')
6. **Credentials:** Name and credentials of all human/ institutional providers of information put up on the site, including dates at which credentials were received
7. **Privacy Policy:** Privacy and data protection policy and system for the processing of personal data, including processing invisible to users, to be clearly defined
8. **Currency:** Clear and regular updating of the site, with date of up-date clearly displayed for each page and/ or item as relevant. (Example: The date on which the last update was done should be mentioned)

9. **Feedback:** User feedback, and appropriate oversight responsibility (such as a named quality compliance officer for each site)
10. **Partnering:** Responsible partnering -all efforts should be made to ensure that partnering or linking to other websites is undertaken only with trustworthy individuals and organisations who themselves comply with relevant codes of good practice
11. **Editorial policy:** Clear statement describing what procedure was used for selection of content.
12. **Searchability/ Findability:** Use of search engine to aid users in searching for information within website
13. **Guidelines:** Tips on searching and finding information within the site
14. **Readability:** Tools to aid user with readability of technical jargon and information (like glossary, link to medical dictionaries, etc. Some websites may highlight particular words which will be links to their meanings on medical dictionaries)
15. **Accessibility:** Enabling website access by physically challenged patients.
16. **Susceptibility:** Information or links to information on susceptibility factors for breast cancer (diet, hormonal factors, environmental factors, lack of exercise).
17. **Severity:** Information or links to information on different stages and grades of breast cancer.

18. **Cues to action:** Phone numbers for patients to call, links to other breast cancer websites, links to other magazines, etc.
19. **Treatment:** Information on any treatment methods/treatment options/treatment advice for breast cancer.
20. **HON Code:** Does website carry the HON Code Seal?
21. **Information:** Website carries any breast cancer information
22. **Commerciality:** Website carries commercial information (advertisements, memorabilia, etc.)

CHAPTER 6: CONTENT ANALYSIS

In order to carry out the comparative evaluation of corporate versus not-for-profit websites carrying cancer-related information, it has been decided to employ a content analysis of the various websites chosen. The quantitative content analysis would be performed on the varying attributes of all the content that is displayed on the home page of these websites and also its different sections and links.

In the backdrop of this study, the definition of content analysis that would be most apt to adapt would be that of Kerlinger (1968): “Content analysis is a method of studying and analyzing communication in a systematic, objective, and quantitative manner for the purpose of measuring variables.”

In this case the seven variables that have been previously stated in this paper would amalgamate to form one large pervasive variable called “quality”. In other words, the overall quality of the websites will be measured by integrating the measurements of their scores as regards to credibility, content, disclosure, links, design, interactivity and caveats.

The use of content analysis in this project is based on one of the principles as stated by Wimmer (2000), where he states that content analysis is used to “test the hypotheses of message characteristics”. In other words, he states that “a number of analyses attempt to relate certain characteristics of the source of a given body of message content to the characteristics of the messages that are

produced.” In this context, we are attempting to relate the quality of content to the monetary agenda of the organization that owns the website.

However, the final analysis is aimed at helping the reader determine if any websites among these two groups of cancer websites are of overall better quality than the other, and if so, what are the attributes that make one better than the other. This would also lead to a set of recommendations for the group with the lesser quality, as to the attributes they need to improvise on in order to better the quality of their own websites.

A widely used method in the realm of health communication, content analysis has been used in a wide-ranging array of ways to assess different outcome measures. In one such study, Im and Chee (2003) examined the various feminist issues in an e-mail group discussion among cancer patients. In order to carry out this project, the researchers recruited participants from Internet Cancer Support Groups (ISCG’s) that were identified through search engines Yahoo and AltaVista. After this, the researchers posted nine different topics for discussion on a separate website created for purpose of this project and asked the participants to carry out the discussion solely through e-mail. What followed was a thorough content analysis of all the messages that were exchanged between the study participants which resulted in a discovery that there were gender and ethnic differences in specific cancer pain experiences like “sharp pain”, “hot burn pain” and “punishing pain.”

Another study by White & Dorman (2000) examined the online support for caregivers by carrying out an analysis of an Internet Alzheimer mailing group. After identifying that Alzheimer's support groups focus primarily on the caregiver than the patient, the researchers examine the content and themes of 532 messages posted on a public Alzheimer mail group during 20 days of 1998. They divide the content at hand into eight quantifiable attributes such as: Information giving/ information seeking, encouragement/support, personal experience, personal opinion, prayer, thanks, humor and miscellaneous and carry out a simple quantitative study by measuring the number and range of list postings by month and then, the number of messages posted by respective list members. This study is a solid example of how easily quantifiable attributes result in smoothly achieving set research goals.

6.1: ADVANTAGES AND DISADVANTAGES OF CONTENT ANALYSIS FOR THIS PROJECT

Very clearly, content analysis has aided me in carrying out a clear compartmentalization of all the content that has been gathered from websites into quantifiable attributes. This has not only further help in carrying out a simple statistical analysis but also eventually helps me answer the research question that has been set at the beginning of this research study. In view of the fact that the medium under consideration is the Internet, it is well known that most major websites update their pages regularly, and hence, content analysis will allow me to take repeated measures of the same websites over a period of time – that is, even after this particular thesis project has been completed. I perceive this as a major advantage since this will allow me to study the changes in the overall quality of cancer-related e-information over a larger period of time. Finally, this method will prove to be a cost-effective yet apt research method since most of the samples will be chosen from the Internet and the analysis conducted on SPSS, the software that is easily procurable in this time and day.

However, the limitations of employing this method lie in the fact that the results of this content analysis will be limited to the framework of the categories and the definitions used in this analysis. In due course, different health communication researchers might use other definitions of the same categories, maybe with differing results and analysis. Thus, content analysis is not an all-pervasive method for this kind of research. Finally, the boundaries of this content analysis have been restricted to about 40 websites in only the ‘cancer’ health

domain. Thus, answers pertaining to the overall quality of websites in other, equally or more dangerous illnesses will remain unanswered.

Reliability: Despite the limitations that have been stated above, it is clear that this content analysis will meet the “reliability” condition due to a number of reasons. First, the category boundaries have already been well defined by a participatory conglomeration of over 60 delegates from governments, industry, academia, and medical and consumer associations including the HON Foundation. Secondly, the categories in the instrument are such that the extraneous variables (like frequent site update operations) will be controlled over a period of time. Even if the sites are updated frequently, these are certain website characteristics that won’t be too dynamic in their change.

Validity: There have not been any studies to date that have used the European Union online health information quality recommendations for conducting their research. This can in part be attributed to the very limited amount of research conducted in this area. Neither have any studies citing this particular instrument as part of their project, mentioned any validity measures. It will thus be extremely challenging to determine the validity of this particular instrument, though in the case of this project, we will assume that it will measure what it sets out to measure – the overall quality of cancer-related websites.

6.2: LIST OF RESEARCH QUESTIONS (RQS)

RQ1: Which of the two groups of websites – commercial and nonprofit – has better quality of online breast cancer information?

RQ 2: What is the nature and extent of the difference between each of the quality parameters when applied to both the website groups individually?

RQ 3: Which of two groups of websites has a stronger presence of information related to individual constructs from the Health Belief Model?

RQ 4: Which of the two groups of websites has a stronger presence of: a) any kind of breast cancer information, and b) commercial information?

RQ 5: Among websites that have any kind of breast information, how many of their web pages provide information on the source and credentials of the source of the respective information?

RQ 6: Is there a difference between the two groups of websites in being certified by the HON Foundation code of conducting and carrying their seal? If so, what is the extent of this difference?

Though it is apparent from my statement of research that I aim to make a comparative analysis of two organizational groups of cancer-related websites, there are a few more hidden questions that I hope this research will help me answer.

Firstly, I want to examine and discover if there exists a relationship between the monetary agendas of organizations and the quality of the content on their websites; and if so, whether the effect poses positive or negative implications. This will lead me to explore deeper research agendas like determining whether a certain standard needs to be created in order to minimize the quality of health information on the internet, given the fact that the content on health websites can be crucial to people's lives and is a matter of great concern and urgency. Secondly, there has always been an urge to discover if the type of organization reflects on the accuracy of content carried on this website. However, the largest must be stated last. For, above all this, I have always been intrigued by the process of evaluation of website content, given the quantum amount of information available. Through this study, I hope to answer my own question, if the quality of website content can actually be quantified and later evaluated if categories are well defined and the quantifying instrument is well aided to suit the purpose of the content under scrutiny.

CHAPTER 7: RESULTS AND FINDINGS

Forms and tables were created on Microsoft Access software in order to key in the codes on each of the 24 quality parameters for 627 web pages across the 20 websites. There was a total of 395 commercial web pages and 232 non-profit web pages. Upon completion of the coding, the data were then exported to Microsoft Excel to create a final dataset.

Summary Statistics:

Table 1: Number of web pages in Commercial and Nonprofit websites

Commercial websites	First layer Web pages	Nonprofit Websites	First layer Web pages
A	57	K	32
B	24	L	19
C	47	M	15
D	30	N	28
E	17	O	11
F	44	P	47
G	83	Q	18
H	15	R	19
I	24	S	11
J	54	T	32
Total	395	Total	232

The analyses for this project were conducted on SPSS Software V.13.

Four sets of analyses were performed on the data:

7.1 FACTOR ANALYSIS

Factor analysis followed by Varimax Rotation was used to reduce the dimensionality of the initial item pool (Gorsuch, 1983). Items with eigenvalues of 1.0 or higher and factor loadings of 0.20 were retained (Nunnally & Bernstein, 1993). The four concepts that were operationalized, Transparency & Honesty, Authority, Accountability and Accessibility, yielded five factors. Specifically, Transparency and Honesty yielded two factors (Name & Address, Purpose, Target Audience; Transparency) that accounted for 82.8 % of the total variance. Authority yielded one factor comprising Sources & Credentials, accounting for 78.35% of the total variance. Subsequently, the Accountability construct (comprising Feedback, Partnering, Editorial Policy) was analyzed. After excluding feedback due to low factor loading, this yielded one factor comprising Partnering and Editorial policy that accounted for 39.10% of the total variance. Finally, the Accessibility construct was tested using the following variables: Guidelines, Searchability/Findability, Readability, and Accessibility. However the Accessibility item had to be excluded even before the matrix was computed, due to zero variance. This construct finally yielded one factor (comprising Guidelines, Searchability/Findability,

and Readability), eventually accounting for 38.51% of the total variance. Indices were created for each of these four new factors by computing summations of the respective items that constructed them.

The final set of EU Foundation's quality parameters thus comprised a total of 9 items: TRANSHON (comprising Name & Address, Purpose, Target Audience), AUTHORITY (comprising Sources and Credentials), ACCOUNTABILITY (comprising Partnering & Editorial policy), MAINACCESS (comprising Guidelines, Searchability/Findability and Readability), Transparency, Currency, Feedback, Accessibility and Privacy Policy. The other parameters on which our analyses were conducted were: HBM Constructs: Severity, Susceptibility, Fear, and other quality constructs: Information, Commerciality, Treatment and HON Code.

7.2 RELIABILITIES

Second, a reliability analysis was conducted using Chronbach's α . The reliabilities that resulted included TRANSHON = 0.668, Authority = 0.688, Accountability = 0.052 and MainAccess = 0.045. The reliability scores are displayed in tables 2-5.

Table 2: Parameters of TRANSHON (Transparency and Honesty) - Varimax Rotational Factorial Structure

<i>Facet</i>	<i>Item</i>	<i>Corrected α if Overall Item Total item Chronbach Correlation deleted</i>	<i>α</i>
TRANSHON	Name Address	& 0.539	0.536
	Purpose	0.697	0.418
	Target Audience	0.689	0.455
	Transparency	-0.001	0.837

Table 3: Parameters of Authority: Varimax Rotational Factorial Structure

<i>Facet</i>	<i>Item</i>	<i>Corrected Item Correlation</i>	<i>Alpha Total if item deleted</i>	<i>Overall Chronbach α</i>
Authority	Sources	0.567	a	0.688
	Credentials	0.567	a	

a. The value is negative due to a negative covariance among items.

Table 4: Parameters of ACCOUNTABILITY - Varimax Rotational Factorial Structure

<i>Facet</i>	<i>Item</i>	<i>Corrected Item Correlation</i>	<i>Alpha Total if item deleted</i>	<i>Overall Chronbach A</i>
ACCOUNTABILITY	Feedback	-0.023	0.276	0.052
	Partnering	0.08	-0.038 (a)	
	Editorial Policy	0.051	-0.022 (a)	

a. The value is negative due to a negative average covariance among items

Table 5: Parameters of Accessibility - Varimax Rotational Factorial Structure

<i>Facet</i>	<i>Item</i>	<i>Corrected Item Correlation</i>	<i>Alpha Total if item deleted</i>	<i>Overall Chronbach α</i>
MAINACCESS	Searchability/Findability	0.034	0.253	0.045
	Guidelines	0.029	0.039	
	Readability	0.082	0.001	

7.3 T-TESTS

In order to evaluate the differences between the two categories in the context of specific variables, we performed four different sets of Independent T-test analysis. The results of the analyses are summarized in Tables 2-5. First, we ran a t-test to evaluate the overall quality score of both the groups. The overall quality (QUALITY) is calculated by adding values of all the nine HON Foundation items to get a cumulative score. The results of this computation are summarized in the table below:

Table 6: Group Differences for Overall Quality scores between Commercial websites and non-profit websites

Quality Measure	<u>Commercial</u>		<u>Non-profit</u>		t
	M	SD	M	SD	
Overall Quality	1.33	1.08	2.06	1.97	-6.01***

*p < 0.05, *** p ≤ 0.001

Significant differences were found between the overall quality scores for commercial and non-profit websites. Non-profit websites had a significantly higher mean (2.06) as compared to commercial websites (1.33).

Next, we performed an Independent T-test analysis to test differences for each of the 9 HON Foundation Quality parameters between the two groups of websites. The results for this test are summarized in Table 7:

Table 7: Group Differences for individual quality parameters between Commercial websites and non-profit websites

HON Foundation Quality Measures	<u>Commercial</u>		<u>Non-profit</u>		t (625)
	M	SD	M	SD	
Transparency & Honesty	0.10	0.40	0.86	1.18	-11.74***
Authority	0.23	0.52	0.21	0.57	0.37
Accountability	0.02	0.12	0.03	0.18	-0.87
Mainaccess	0.53	0.53	0.68	0.50	-3.40***
Transparency	0.12	0.32	0.05	0.21	3.00*
Privacy Policy	0.05	0.22	0.03	0.17	1.22

Currency	0.24	0.43	0.18	0.39	1.82
Feedback	0.04	0.18	0.03	0.16	0.658
Accessibility	0.00 ^a	0.00	0.00 ^a	0.00	-

*p < 0.05, *** p ≤ 0.001; a. t cannot be computed because the standard variations of both groups are zero.

Significant differences were found between commercial and non-profit websites for three of the nine quality parameters. Non-profit websites scored a higher mean on Transparency and Honesty (0.86) as compared to commercial websites (0.10).

Additionally, it was found that non-profit websites scored significantly higher in the MAINACCESS score (comprising searchability /findability, readability and guidelines) with a score of 0.68 as compared to commercial websites (0.53). Commercial websites, on the other hand, recorded marginally higher mean values for Transparency (revealing sources of funding) with a score of 0.12 as compared to non-profit websites 0.05. No significant differences were found between the two groups for the remaining six quality parameters. It must be observed though, for the record, that commercial websites scored marginally higher on the 'currency' of information score, than did non-profit websites. It is also critical to note from the above table, that not one of the 627 web pages

across both groups of websites was specifically enabled for physically challenged people.

Next, we tested the differences between the two groups of websites for types of information in line with the Health Belief Model (HBM) constructs.

The results of this analysis are summarized in Table 8:

Table 8: Group Differences for Health Belief Model (HBM) construct scores between Commercial websites and non-profit websites

Quality Measure	<u>Commercial</u>		<u>Non-profit</u>		t (625)
	M	SD	M	SD	
Susceptibility	0.07	0.25	0.05	0.22	0.713
Severity	0.03	0.17	0.02	0.26	0.84
Cues to action	0.16	0.37	0.26	0.44	-2.95*

*p < 0.05, *** p ≤ 0.001

There were no significant differences in the presence of susceptibility or severity information between commercial and non-profit websites. What was significantly different though, was the presence of ‘Cues to Action’ between the

two groups, where non-profit websites scored significantly higher means (0.26) than commercial websites (0.16).

Table 9: Group Differences for Information, Commerciality and Treatment scores between Commercial websites and non-profit websites

Quality Measure	<u>Commercial</u>		<u>Non-profit</u>		t (625)
	M	SD	M	SD	
Information	0.42	0.49	0.47	0.50	-1.33
Commerciality	0.58	0.49	0.07	0.26	14.55***

*p < 0.05, *** p ≤ 0.001

Out of the 627 websites analyzed, commercial websites were found to have significantly higher commercial content (0.58) compared to nonprofit websites. Contrastingly, nonprofit websites had more breast cancer information than commercial websites but the difference was insignificant.

7.4 CROSSTABULATIONS

Chi-square tests of independence were performed to examine the relationships of information with authority, HBM Constructs and Treatment. The relationship between information and authority (Table 10) was significant, $\chi^2 (2, N=627) = 13.52, P \leq 0.001$

Though commercial websites had a relatively higher presence of sources of information stated, nonprofit websites had a higher presence of both, sources and credentials, stated.

Table 10: Web pages with any “information” on breast cancer meeting the parameter of “authority”

Authority	<u>Commercial</u>	<u>Nonprofit</u>
	% within website	% within website
0.00	82.0	87.1
1.00	13.2	4.7
2.00	4.8	8.2
χ^2	13.52***	

* $p < 0.05$, *** $p \leq 0.001$

Table 11: Web pages with any “information” on breast cancer providing information on ‘susceptibility’ factors

		<u>Commercial</u>	<u>Nonprofit</u>
Susceptibility		% within website	% within website
Not Present		93.4	94.8
Present		6.6	5.2
χ^2	0.51		

* $p < 0.05$, *** $p \leq 0.001$

The relationship between information and Susceptibility (Table 10) was not significant, $\chi^2 (1, N=627) = 0.51$. It was found that commercial websites reflected a stronger presence of information of information on susceptibility factors (risk factors of breast cancer (14.0), than nonprofit websites (10.14).

Table 12: Web pages with any breast cancer ‘information’ that has ‘severity’ information

	<u>Commercial</u>	<u>Nonprofit</u>
Severity	% within Information	% within Information
Not Present	93.9	96.3
Present	6.1	3.7
χ^2	11.37***	0.04*

* $p < 0.05$, *** $p \leq 0.001$

The relationship between information and severity (Table 12) was significant, χ^2 (2, N=395) = 11.37, $P \leq 0.001$ for commercial websites; χ^2 (2, N=232) = 0.04, $P < 0.05$ for nonprofit websites. Commercial websites, it was found, had a marginally higher % of information about severity (stages and grades) of breast cancer with 6.1% than nonprofit websites that scored 3.7.

Table 13: Web pages with breast cancer ‘information’ that provide information on ‘cues to action’

	<u>Commercial</u>	<u>Nonprofit</u>
Cues To Action	% within Information	% within Information
Not Present	84.8	71.6
Present	15.2	28.4
χ^2	0.19	0.71

* $p < 0.05$, *** $p \leq 0.001$

The relationship between information and cues to action (Table 13) was not significant, $\chi^2 (2, N=395) = 0.19$ for commercial websites; $\chi^2 (2, N=232) = 0.71$ for nonprofit websites. Nonprofit websites were found to carry more information about cues to action – encouraging/motivating audiences to make phone calls for motivation therapy, read magazines of breast cancer, etc. – with 28.4% than commercial websites with 15.2%.

Table 14: Web pages with breast cancer ‘information’ providing treatment information

	<u>Commercial</u>	<u>Nonprofit</u>
Treatment	% within Information	% within Information
0.00	82.0	87.1
1.00	13.2	4.7
χ^2	22.34***	17.54***

* $p < 0.05$, *** $p \leq 0.001$

The relationship between information and authority (Table 14) was significant, χ^2 (2, N=395) = 22.34, $P \leq 0.001$ for commercial websites; χ^2 (2, N=232) = 17.54, $P \leq 0.001$ for nonprofit websites. Commercial websites were found providing more information on treatment options for breast cancer (13.2%) than non-profit websites (4.7%).

CHAPTER 8: DISCUSSION

8.1 OVERVIEW

This chapter will conclude our comparative analytical evaluation of the quality of online breast cancer information in commercial and breast cancer websites. In doing so, we will examine the performance of the two groups of websites with respect to their performance against different quality parameters and constructs. Following that, we will juxtapose the same examination in the backdrop of the theoretical framework of the Health Belief Model. Additionally, this study faced a number of critical challenges in its analyses phase, especially with regard to deciding on a unit of analysis and the nature of tests to be conducted in order to achieve as close to an unbiased result as possible. Considering the extremely limited amount of previous research conducted in this area, there was little reference material. The above analyses offer us, what we conceived to be the strongest combination of possible tests for the required examination, but there are abundant opportunities for improvement. Thus, we will also identify areas where future research can improvise upon any addition this study would have made to the existing body of knowledge in online breast cancer communication— improvisations as much in terms of methodology and statistical analysis as in terms of the basic research idea. We will close the discussion by making recommendations for future work that can be carried out in this very fascinating field of consumer health informatics.

8.2 ANALYSES OF THE RESULTS

RQ: Which of the two groups of websites – commercial and nonprofit – have better quality of online breast cancer information?

The evaluation of breast cancer information quality among 627 web pages across commercial and breast cancer websites was primarily based upon the health information quality criterion set by the HON Foundation, and health behavior constructs as posited by the Health Belief Model. This resulted in the information being evaluated against a total of 15 quality parameters.

Considering this, it can be thus be deduced that the maximum possible mean for the overall quality score for each of the groups is 15. Though nonprofit websites had a significantly higher overall quality score of 2.12 as compared to commercial websites with a score of 1.32, it is evident that neither of the two groups is particularly strong on meeting the quality parameters; in fact the scores are abysmally low. What is of great concern is the fact that these websites were listed as the 10 most popular websites in their respective categories by Google at the time when the search was ran and saved in May 2005.

RQ: Which of the two groups possess a) a stronger presence of any kind of breast cancer information and b) commercial information?

Though both the groups did not differ significantly in terms of individual web pages with any information on breast cancer, the commercial websites had far more commercial content than nonprofit websites.

This commercial content has been identified as primarily being in the form of banner advertisements for credit cards, travel deals, etc. Many commercial websites were also found to solicit sales for breast cancer memorabilia like bracelets, anklets, sweatshirts, etc. The implications of heavy advertising on 'breast cancer websites' merits examination way beyond what has currently been accorded. This is because such banners possess the potential to distract the patient and/or his/her relatives/friends from important information-seeking behavior on breast cancer and can divert their attention towards areas that might not necessarily be of health-related utility to them.

This is in no way meant to suggest health-related websites must be completely devoid of banners – that might be challenging some basic economic precincts of commercial breast cancer informatics. However, advertising could be regulated in a way that the best interests of the information-seeking consumer are kept in mind and also probably, in ways that encourage recommended health behaviors, which, for breast cancer patients, might be attending screening sessions. The potential of banner advertising to create large-scale product awareness in a short span of time is well known and it is this potential that needs to be tapped to encourage health behaviors. It might also be interesting to

conduct detailed examinations of what exactly might be the recommended percentage of banner advertising on health-related websites. This kind of examination might well help in discovering the threshold of the consumer with regards to the point when the banners start intruding into his information-seeking behavior.

RQ: What is the nature and extent of the difference between each of the quality parameters when applied to both the website groups individually?

Low quality of online health information has for long been the pet peeve of health communication researchers. It is unfortunate that this study reinforces and acknowledges this conception, lest the very objective be lost. The findings regarding the performance of these two groups of websites in the context of specific quality parameters provide some interesting insights into the nature and extent of differences between commercial and nonprofit websites in the context of specific quality parameters. Interestingly, neither of the two groups showed a consistently high or low score across all parameters. While the nonprofit websites had a stronger presence of certain parameters, the commercial websites scored better on others.

Specifically, the findings of the independent t-test analysis showed that nonprofit websites had a much higher mean on the Transparency-Honesty score. What this translates to be that there were more nonprofit websites specifying their name/address, purpose and the target audience, than commercial websites. Though previous research has not been conducted in this specific domain, it

might be worthwhile to evolve a rationale for this. The principal guiding factor for commercial breast cancer websites is obviously more economics than service. Hence, specifying the target audience might well potentially mean forfeiting a certain (though possibly small) percentage of the remaining and thus a possible source of future revenue. Though the exact amount of incidence of the purpose/objective being stated by commercial websites has not been ascertained, it is well known that nonprofits are more purpose-oriented than commercials that are revenue-oriented. Similar caveats hold true for the name/address variable.

Nonprofit websites also scored significantly higher than commercial breast cancer websites in terms of providing a more easily readable experience for their audiences. By this we mean, that nonprofit websites had a higher presence of intra-site search engines, web pages devoted exclusively to specifying guidelines for content and so too, providing readability features like online dictionaries that would make their information more easily readable. Though some commercial websites too met these parameters, the presence was not as prominent as it was in nonprofit websites. There is anecdotal evidence from studies in the area of real estate about commercial websites ranking lower than nonprofit websites in constructs similar to these, and the explanations again point towards the heavy stress on profit and revenue than enhancing their reader experience. It might be interesting to explore if commercial websites are, in fact, monetarily more robust than nonprofits in order to provide a better reader experience. Additionally, commercial websites might have access to better quality web designers whose insights and expertise can be incorporated into their web pages. Whether

providing an enhanced reader experience figures in the list of organizational priorities for commercial websites, and if they do, how high, would eventually determine the future course of commercial websites bettering their performance in light of these parameters. There are also other apparent incentives in this endeavor for commercial websites in that, providing a better integrated website navigation system can be tied to their advertising model in order to maximize revenue. Of course, nonprofit websites too have a long way to go in providing their readers with a better experience through more powerful site navigation and it must not be lost on us that our analysis so far has only been conducted in light of their comparison against commercial websites.

The lower mean of nonprofit websites in stating their sources of funding as compared to commercial websites is as much intriguing as it is interesting. This is because nonprofit websites are perceived to clearly state their sources of funding and donor agencies. What could probably explain this finding is the fact that the nonprofit websites researched upon varied in size and this could have probably reflected on the funding. Like most other parameters, it is not one of the two, but both groups that need to provide clarity about their sources of funding. Providing this information upfront extinguishes the possibility of any doubt in the mind of the audience as to who is it that is financially supporting the display of information that is being accessed. This provides a certain level of legitimacy to the informational environment that is being experienced by the audience.

RQ: among websites that have any kind of breast information, how many of their web pages provide information on the source and credentials of the source of the respective information?

It must be observed that though commercial websites show a greater presence of sources for all the information on their pages, nonprofits have a greater percentage of instances with both sources and credentials. It is certainly pertinent to think if this reflects on the concern for informational authority respectively by the two groups. While commercial websites might contend with only stating the source to provide legitimacy to their information, nonprofit websites insist on also providing the credentials for their sources of information.

The importance of treatment-related information on health websites has often been pointed out in previous research and nonprofits score a significant point over commercial websites. One reason for this might be that commercial websites might deliberately consider desisting from providing treatment options in order to avoid running the risk of a wrong prescription. However, it can also be argued that commercial breast cancer websites that are affiliated to say, a breast cancer organization or a pharmaceutical company can well detail treatment options as a way to furnish their monetary interests.

One of the more significant findings from the analysis is that none of the 627 web pages analyzed were made accessible to physically handicapped people. Though it can be argued that the percentage of breast cancer patients who are also physically challenged might not be all that great, it does not warrant that this entire spectrum be left out of consideration while designing a website.

Making a website accessible to physically challenged people is one of the basic premises in web design, especially for health websites, and it is a matter of great concern that this has not even been addressed once among the ten most popular websites in each category.

8.2 LIMITATIONS

In this study we attempted to carry out a comparative evaluation of breast cancer information quality in commercial and nonprofits websites. There can be several limitations to such a study.

First, the 20 most popular websites analyzed for this study were chosen only at a specific point of time and not over a longer period of time. Thus, the results of this study might be very specific to this cohort of websites and not easily generalizable to the larger population.

Next, using Google's advanced search facility, which gives results based on their page-rank index, we chose the websites. Though the selection of the websites for this study has been fair, the results offered by Google change at a rapid pace, at times daily, and the 20 most popular sites selected for the study in May 2005 might not be the same results that come up today. This can also partly be attributed to the rapid proliferation of health websites and this is one challenge, I content, most researchers in this area might encounter.

Speaking of statistical limitations, it is identified that the use of factor analysis for analyzing dichotomous variables may face criticism. This is especially because some of the items that were force-factor analyzed based on preconceived categories had extremely low reliability scores. Also, we had to test some of the times that had almost zero variance, hence giving rise to the possibility that the factor analysis results could be artificial. However, given the large number of variables it was deemed necessary to use factor analyze as a

data reduction tool and given that some of the factors found had high reliabilities, the use of this method was not completely unwarranted. There is a potential to use the websites, than the web page, as a unit of analysis to tide over this issue and this potential must be explored in further studies with a greater degree of statistical sophistication.

8.3 THEORETICAL IMPLICATIONS

Given that the Health Belief Model's roots are based in the Stimulus – Response (S-R) Theory, it is tempting to suggest that the three constructs of the HBM that have been tested in this study have been treated as 'reinforcements'. In other words, we have attempted to study the presence of susceptibility, severity and cues to action information on breast cancer websites since we posit that that these might affect physiological drives that finally impact behavior. Of course, the lack of mentalistic concepts such as reasoning or thinking that characterized the S-R theory will not operate here, since information processing includes interpreting the read information and forming logical connections to set it up against one's own circumstances.

In this context, we must revert back to the findings to note that on any given web page that had some sort of breast cancer information, commercial websites addressed susceptibility and severity constructs from the health belief model constructs to a greater extent than nonprofit websites did. In light of this, it might not be unreasonable to argue that providing information on the risk factors for cancer and so too the stages and grades of cancer depend to a certain extent, on manipulating the fear component in the audience's minds. Though there might be contrasting points of view, it can definitely be argued that the fear factor has always been used as a potent persuasive tool and might well be an indirect revenue model for commercial websites. On the other hand, nonprofit websites have proven to be stronger on the 'cues to action' construct (an HBM construct

which has been adequately criticized for lack of clarity but is of ample use here) that is definitely closer to motivating towards intended behavior. For example, some of the websites that were analyzed had ‘trigger points,’ where audiences could first read the information and then call up hotlines for more information on already present or new information. This section also included other trigger points such as suggesting breast cancer magazines to be read, attending information sessions by breast cancer survivors, etc. The most interesting aspect of our analysis however, has been the observation of a dynamic interplay between these three constructs.

Hereon, the moot point with implications for theory is the notable significance of Health Belief Model constructs in the context of online breast cancer information. Traditionalists might readily agree on the old Health Belief Model paradigm where in the five HBM constructs finally lead to performance of the end behavior, which is seeking online breast cancer information. However, it might also be useful to provide a contrasting alternative for employment of this model in this area.

With increasing knowledge and use of health-related websites, it is also possible that the patient and/or her family members might first evaluate the benefits and barriers of online health information. In an ideal scenario, the presence of more perceived benefits than barriers might lead to their seeking information online. Hereon, it is very possible that the principle motive for accessing online health information is to seek relevant, good quality information about the susceptibility and severity factors of breast cancer. Given that within

breast cancer, treatment options are most frequently searched topics (Ziebland, et al.) their online behavior might lead to seeking information on treatment options, medications and tips for healthy behavior. The quality and type of information being accessed here might eventually result in a healthy behavior – like reducing alcohol use and increasing exercises.

To summarize, the HBM constructs can thus be operated upon in two ways: 1) online health information seeking as an end behavior in itself and 2) online health information sources as Reinforcement tools to motivate the end user towards other healthy behaviors. It will be interesting to see the implications of this for future research and these will be discussed later on in the “Recommendations for Future Research” section.

One of the arguments for the use of HBM in this particular study might be that only three out of the five constructs were measured. This was primarily because “benefits” and “barriers” cannot be operationalized and converted into measures with the same flexibility as constructs like susceptibility, severity and cues to action, can. By this we mean, perceptions of barriers and benefits are mostly psychological or maybe even psychosocial and thus cannot be treated as quality parameters of a website or be measured as a type of information. The perceived barriers and benefits can certainly be examined with the help of experiment research using human subjects and this is one of the directions this research has the potential to take in the future.

8.4 PRACTICAL IMPLICATIONS

Our analyses of commercial and nonprofit breast cancer websites adequately support existing literature that comments about the dissatisfying quality of online health information. In this light, it can be said that this study has significant implications for breast cancer website creation, development and maintenance in the years to come.

It has now been proved in more than just one instance that 'quality of information' is not an area that health website creators have strongly considered in the process of deciding on the site architecture and content. While commercial websites are mainly being driven by their monetary interests in making websites that focus comparatively more on content that is due to drive sales of their respective products (pharmaceuticals, memorabilia, etc.), nonprofits have probably suffered in creating an efficient navigation and content system due to financial hiccups. Either way, little effort has been made to ensure that the information, which in this case can be extremely crucial, conforms to established quality parameters in good measure.

From this, it also follows that though many organizations have so far established their own codes of conduct for online health information and parameters for health information quality there has not been a single, major effort to create a standardizing mechanism for such content. Indeed, it is unfortunate that despite the criticality of such information and also especially the criticality of breast cancer in women's health, little initiative has been taken to create a

common thread of rules which will push website creators to be more conscientious in their content. The difference is between either having a set of rules that website owners may/may not want to consider or establishing one that they 'have' to consider in order to be hosted on the Internet. Evidently, the former does not seem to be reaping dividends and hence the latter option might be the path to take if we are to ensure that breast cancer patients and their families get information that is reliable and recent.

Outside of creating a common set of parameters for all breast cancer websites in order to get them to a higher quality level, what must also be explored is the option of using the concept of RSS/XML. These feeds that have so far been used by users to get customized news content on their preferred websites can possibly be mechanized to an algorithm comprising quality parameters where the search engine only offers results of website that conform to those parameters. For example, we can create an algorithm by which a search engine first evaluates a website based on the HON Foundation parameters and then, comes up with results for the user. Such an inbuilt operation will facilitate in reducing 'information scanning' time for the user who can now readily step up to the 'information seeking' level. However, which set of quality parameters need to be adopted for such an approach and their implications on user behavior will have to be further explored to greater detail. It will take a mammoth effort to create such an overarching standardizing mechanism but once employed, the results will be there for all to see. It is only a matter of recognizing the criticality of this issue and lending it the importance it deserves.

8.5 RECOMMENDATIONS FOR FUTURE RESEARCH

Based on the findings and analyses of this study and previous literature, it can be safely said that there are still many fresh dimensions and areas in e-health that need to be examined. The encouraging aspect is that the multidisciplinary nature of e-health allows us to employ perspectives from a variety of other fields of knowledge like information sciences, communications, behavior sciences, economics and sociology. Additionally, the emergence of 'infonomics' as a new field of study is a promising prospect. Governments and academic institutions too are rising up to the fact that informatics is impacting on healthcare delivery in ways more rapid than one imagined to be – and at the very core of this is the internet, a medium innovation that has been diffused and adopted at a rate faster than traditional media. These are reasons enough to make us believe that the amount of scholarly work conducted in this area will increase multifold in the near future.

Firstly, it must be noted that the attempt of this study has been to strictly answer the six research questions that were outlined at the very onset. The dataset used for this study can be used to explore and examine a number of relationships that might possibly exist between and within variables and groups of variables, using different permutations and combinations. For example, it might be interesting to examine whether the sites that do define their editorial policy, back it up with sufficient authorship information. Secondly, this study is restricted to evaluating if or not the websites meet certain quality criteria. An

interesting extension of this research might be to perform a content analysis and evaluate the propriety and authenticity of the information, possibly leading to the creation of new standards for online health information. It might also be useful to lay additional stress on performing advancing statistical tests in order to obtain more information about quality than can be traditionally be obtained. A useful lesson learned from this study was the efficient use of factor analysis as a data reduction tool, especially given the huge number of variables (22) that were present at the commencement of the study.

One of the limitations of this study is the fact that two very important HBM constructs, Perceived Benefits and Perceived Barriers, were not tested. This can partly be attributed to the fact that testing and measurement of these constructs would involve the use of human subjects and given the logistical limitations to this study, it was not possible. However, it would be interesting to devote scholarly work into analyzing the exact interplay of the Health Belief Model constructs and online information seeking behavior. Considering that our findings revealed a greater stress on susceptibility and severity by commercial websites and on cues-to-action by nonprofit websites, there is now a possibility to explore the creation of new models predicting how audiences can possibly streamline their searches depending upon the type of information they are looking for. Additionally, it can also be explored whether the kinds of information being accessed differs between the roles than the user plays. For example, while families/friends/ relatives of breast cancer patients might be seeking a particular kind of information, the patients themselves might be looking for other kinds.

From this it follows, that testing traditional health behavior theories in the backdrop of e-health would not only help these theories evolve further, but also provide fresh perspectives to e-health experts.

Future research can also build upon this study by examining if online health information has any direct or indirect impact on final behavioral outcomes. Conducting such studies will lead to the identification of specific content categories that have highest impact and these categories can then be used to customize information content for specific audiences.

Internet health information is now coming to feature in a variety of prominent initiatives by the U.S. federal government. Foremost is this the Office of Disease Prevention and Health Promotion (ODPHP) study in this area that also figures prominently in the Healthy People 2010 plan. Given that the U.S. government is active in international development efforts through organizations like United States Agency for International Development (USAID), World Health Organization (WHO) and the United Nations (UN), it will be interesting to see how this domestic research effort is scaled up to the global health scene.

Given the attention that global health issues currently commands, the use of quality health information to alleviate global health concerns is now increasingly gaining prominence. The Journal of Medical Internet Research, one of the most prominent journals in e-health, recently devoted their entire latest edition, Volume 14, to the use of the use of Internet in the global health scenario. Given the immense amount of projects presently underway in the area of

Information Communication Technologies (ICT's) in countries in Africa and Asia, there lies immense of scope for exploring the use of Internet for disseminating high quality health information to the most disadvantaged populations of the world, in ways both scholarly and otherwise.

APPENDIX 1: SAMPLE CODING INSTRUMENT

Yes = 1, No = 0

Name and Address: Name, physical address and electronic address of the person or organization responsible for the site

☐ 1 ☐ 0

Purpose: Purpose and objective of the site

☐ 1 ☐ 0

Target Audience: Target audience clearly defined

☐ 1 ☐ 0

Transparency: Transparency of all sources of funding for site (grants, sponsors, advertisers, non-profit, voluntary assistance).

☐ 1 ☐ 0

Sources: Clear statement of sources for all information provided and date of publication of source (even a single source provided for overall information on the page should be coded as 'yes').

☐ 1 ☐ 0

Credentials: Name and credentials of all human/ institutional providers of information put up on the site, including dates at which credentials were received.

☐ 1 ☐ 0

Privacy Policy: Privacy and data protection policy and system for the processing of personal data, including processing invisible to users, to be clearly defined.

☐ 1 ☐ 0

Currency: Clear and regular updating of the site, with date of up-date clearly displayed for each page and/ or item as relevant. (e.g. The date on which the last update was done should be mentioned)

☐ 1 ☐ 0

Feedback: User feedback, and appropriate oversight responsibility (such as a named quality compliance officer for each site).

☐ 1 ☐ 0

Partnering: Responsible partnering -all efforts should be made to ensure that partnering or linking to other websites is undertaken only with trustworthy individuals and organizations who themselves comply with relevant codes of good practice.

☐ 1 ☐ 0

Editorial policy: Clear statement describing what procedure was used for selection of content.

☐ 1 ☐ 0

Searchability/ Findability: Use of search engine to aid users in searching for information within website

☐ 1 ☐ 0

Guidelines: Tips on searching and finding information within the site

☐ 1 ☐ 0

Readability: Tools to aid user with readability of technical jargon and information (like glossary, link to medical dictionaries, etc. Some websites may highlight particular words which will be links to their meanings on medical dictionaries).

☐ 1 ☐ 0

Accessibility: Enabling website access by physically challenged patients

☐ 1 ☐ 0

Susceptibility: Information or links to information on susceptibility factors for breast cancer (diet, hormonal factors, environmental factors, lack of exercise)

☐ 1 ☐ 0

Severity: Information or links to information on different stages and grades of breast cancer.

☐ 1 ☐ 0

Cues to action: Phone numbers for patients to call, links to other breast cancer websites, links to other magazines, etc.

☐ 1 ☐ 0

Treatment: Information on any treatment methods/treatment options/treatment advice for breast cancer

☐ 1 ☐ 0

HON Code: Does website carry the HON Code Seal?

☐ 1 ☐ 0

Information: Website carries any breast cancer information

☐ 1 ☐ 0

Commerciality: Website carries commercial information (advertisements, memorabilia, etc.)

☐ 1 ☐ 0

WORKS CITED

Aguillo, I. (2000) A new generation of tools for search, recovery and quality evaluation of World Wide Web medical resources. *Online Information Review*, 24, 138.

Bath, P.A. (2003) Development and application of a tool designed to evaluate websites providing information on Alzheimer's disease. *Journal of Information Science*, 29, 279-297.

Berger, A.A. (1995). Essentials of Mass Communication Theory. (1st. Ed), *Media- McLuhan's Theory: Medium is the Message* (pp. 56-58). London: Sage

Bradley, C. J., & Diepgen, T. (2005). The need for online information on the economic consequences of cancer diagnosis, treatment and survivorship. *Journal of Medical Internet Research*, 7(3), e29.

Cooke, A. (1999) Quality of health and medical information on the Internet. *British Journal of Clinical Governance*, 4, 155.

Domaszewicz, A. (2002) Online tools drive health care decisions. *HR Magazine*, 47, 97-101

European Union. eeurope2002: quality criteria for health related websites. Retrieved Nov. 02, 2005, from Health On The Net Foundation Web site: http://www.hon.ch/HONcode/HON_CCE_en.htm.

Eysenbach, G., & Diepgen, T. (1998). Towards quality management of medical information on the Internet. *British Medical Journal*, 317(7171), 1496-1500.

Eysenbach, G., Kohler, C. (2002). How do consumers search for and appraise health information on the World Wide Web? Qualitative study using focus groups, usability tests, and in-depth interviews. *British Medical Journal*, 324, 573-576.

- Funda, M., & Bernstam, V.E. (2002) Breast cancer on the World Wide Web: cross sectional survey of quality of information and popularity of websites. *British Medical Journal*, 324, 577-581.
- Gagliardi, A., Jadad, A.R., (2002) Examination of instruments used to rate quality of health information on the Internet: chronicle of a voyage with an unclear destination. *British Medical Journal*, 324, 569-573.
- Google. Corporate information. Retrieved Nov. 02, 2005, from Google Web site: <http://www.google.com/corporate/tech.html>.
- Gretchen, B., Eliot, M.N. et al (2001). Health information on the Internet: Accessibility, Quality, and Readability in English and Spain. *The Journal of the American Medical Association*, 285, 2612-2621.
- Griffiths, K.M., Christensen, H. (2000) Quality of web based information on treatment of depression: cross sectional survey. *British Medical Journal*, 321, 1511-1515.
- Hargittai, E. (2000). Standing before the portals: non-profit websites in an age of commercial gatekeepers. *The Journal of Policy, Regulation and Strategy for Telecommunications*, 2, 543-550.
- Henderson, K.V., & Cowart, L.B. (2002) Bucking e-commerce trends: A content analysis comparing commercial real estate brokerage and residential real estate websites. *Journal of Corporate Real Estate*, 4, 375-385.
- Im, E. (2003). Feminist Issues in E-mail Group Discussion Among Cancer Patients. *Advances in Nursing Science- Health and the World Wide Web*, 4, 287-298.
- Jorgensen, J. K., Gotzsche, P.C. (2004) Presentation on websites of possible benefits and harms from screening for breast cancer. *British Medical Journal*, 328, 148-154.
- Kerwin, K.E., Madison, J. (2002) The role of Internet in improving healthcare quality. *Journal of Healthcare Management*, 47, 225-237

- Kunst, H., Khan, K.H. (2002). Quality of web-based medical information on stable COPD: comparison of non-commercial and commercial websites. *Health Information and Libraries Journal*, 19, 42-48.
- Losken, A., Burke, R., Elliot, F.L., Carlson, G.W. (2005). Infonomics and Breast Reconstruction: Are patients using the Internet? *Annals of Plastic Surgery*, 54(3), 247-250.
- McKenzie, C.B. (1997). Quality Standards for Health Information on the Internet. *SIM Quarterly*, 3.
- McQuail, D. (2000). McQuail's Mass Communication Theory. (4th. Ed.), *Media Content; issues, concepts and methods of analysis* (pp. 303-331). London: Sage
- Moskowitz, H.R., Itty, B., Asami, R., Fink, C., Flores L. (2002) Arthritis Websites- deconstructing the messages to identify strong versus weak performers. *International Journal of Medical Marketing*, 3, 31-49.
- Rasmussen, T. (2000). Social Theory and Communication Technology. (1st. Ed), *Structuration Theory and New Media – Insights and Problems* (pp. 11-48). Burlington: Ashgate.
- Rimer, B. K., Lyons E. J., Ribisl K. M., and Bowling, J. M. (2005). How new subscribers use cancer-related online mailing lists. *Journal of Medical Internet Research*, 7(3), e32.
- Risk, A., Dzenowagis, J. (2001) Review of Internet Health Information Quality Initiatives, *Journal of Medical Internet Research* 2001, 3, e28.
- Rosenau, P.V. (2003) Performance Evaluations of For-Profit and Nonprofit U.S. Hospitals since 1980. *Nonprofit Management and Leadership*, 13, 401-423

- Salamon, L.M. (1999). The Nature of the Nonprofit Sector (1st.Ed.). *Scope and Structure: The Anatomy of America's Nonprofit Sector* (pp. 23-38). Colorado: Westview
- Shaddock, J., Locke, A., Ward, B., Collins, J. (2002) Cancer Websites Appraised. *Health Expectations*, 5, 179-181.
- Turner, J.W., Grube, J.A., Meyers, J. (2001). Developing an optimal match within online communities: An exploration of CMC support communities and traditional support. *Journal of Communication*, 51, 231-251.
- White, M. Dorman, S. (2000). Online Support for caregivers: Analysis of Internet Alzheimer's mail group. *Computers in Nursing* (4) 168-179
- Witte, K. (1995). Designing Health Messages – Approaches from Communication Theory and Public Health Practice. (1st. Ed), *Fishing for Success* (pp. 145- 166), Thousand Oaks-Sage.
- Wimmer, D. Dominick, R. (2000). Mass Media Research - An Introduction. (6th Ed.). *Content Analysis* (pp. 133-160). Belmont: Wadsworth.
- Ziebland, S., Chapple, A., Dumelow, C., Evans, J., Prinjha, S., Rozmovits, L. (2005). How the Internet affects patients' experiences of cancer: a qualitative study. *British Medical Journal*, 328, 564-569.