

Health Search Tool Evaluation

By

Aseem Kumar

A CAPSTONE

**Presented to the Oregon Health & Science University's
Department of Medical Informatics & Clinical Epidemiology**

**in partial fulfillment of
requirement for the degree of
Masters of Biomedical Informatics**

May 2005

School of Medicine

Oregon Health & Science University

Master of Biomedical Informatics

Certificate of Approval

This is to certify that the Capstone Project of
Aseem Kumar

“Health Search Tool Evaluation”

Has been approved.



Professor in charge of capstone project

May 17, 2005
Date

Table of Contents

Acknowledgements.....	iv
Abstract.....	1
Introduction.....	2
Materials and Methods.....	7
Results and Analysis.....	14
Summary and Conclusions.....	27
References.....	30
Appendix A.....	34
Appendix B.....	36
Appendix C.....	41
Appendix D.....	43

Acknowledgements

I would like to take this opportunity to thank my advisor, Holly Jimison for her vision, guidance and encouragement throughout this endeavor.

Many thanks also to my family and friends for their moral support.

Health Search Tool Evaluation

Abstract:

Objectives: To discover and interpret consumer behavior and satisfaction parameters for online quality health information retrieval using popular search engines.

Design: A combination of both quantitative and qualitative approach using search questionnaires as well as naturalistic observation.

Participants: A total of 66 study participants were observed naturalistically as they searched the web for six health-related questions using Google, Healix and MedSearch. This was followed by evaluation of the top results by three independent health professionals for quality of health information.

Setting: Out-patient clinics building, Oregon Health Sciences University, Portland

Results: Significant differences were observed between overall performances of the three search engines from the study subjects' perspective. A number of qualitative aspects of their search behavior were noted as they performed the searches. The evaluation by the health professionals was significantly different in terms of quality of the web pages retrieved by the search engines.

Conclusions: Further studies are required to deeply analyze and assess consumer search behavior as well as to design strategies to guide them to quality health information on the web.

Introduction:

The Internet continues to expand as a powerful resource for providing health information. According to a study by Eysenbach in 2003, more than 12.5 million health-related searches were conducted on the World Wide Web. [13] Another study on Internet usage reported that about 33% of their patients used Internet to gain information related to health issues. [15] An interesting aspect is that about a third of the patients rely on health searches performed by family and friends on their behalf. [13,15] Hence the audience for health specific information is huge and expanding. Most of these are users seeking fast information, most likely as a second opinion after having consulted their physicians. [4] Although popularity of the Internet amongst health information seekers continues to rise, there have been questions about the quality of this information as well as the validity. In 2002, Gagliardi and Jadad discovered that the rating instruments for health websites are continuously fluctuating. Moreover most of them were either not functional or available but didn't provide information or they weren't validated. [2] Another systematic review assessing the quality of health information concluded that about 70% of studies have reported lack of quality although operational definitions of quality criteria are still required. [14]

Although quality continues to be an issue, other shortcomings of online health information include difficulties with comprehension, access, potential for harm and risk of over consumption. [17] Dr. Berland and colleagues found that online health information is frequently geared towards well-educated users, with all the English-language websites in their study requiring high school reading levels or higher. [12] Studies on Official Pharmacological Web Sites for erectile Dysfunction treatments have

found the content to be superficial, biased, promotional and ineffective in comparing treatment options. [9]

In spite of all its deficiencies and demerits the currency of seeking health information is on the up especially in light of the limited time patients have while consulting their health care providers. And studies have shown that the Web might prove to be an efficacious source if used appropriately. A study for weight loss management programs identified the Internet to be as successful as face-to-face care. [20] The Internet has the potential to provide a robust resource for meeting public health information needs. [12]

Ideally Internet can maximize the efficacy of the limited time patients spend with their physicians. Consumers should be able to learn most of what they need to know from the high-quality web sites before their intended encounter with the physician. According to the Pew Internet & American Life Project, 81% of consumers seeking health information online do so through a search.[15] Many of the users begin search through some popular search engine rather than a prominent health information site.[21]

Diabetes followed by cancer constitutes the most common medical terms amongst health care information seekers. [8] Women, especially pregnant or diabetic are more likely to use the World Wide Web for answers to health-related questions. [19] According to Eysenbach and Kohler consumers looking for health information primarily focus on source, a professional design, scientific touch, language and ease of use when assessing the credibility of a website. [3] Although measures of quality correlate with accuracy of health information, popularity is decided by the type rather than the content of a website. [6]

Research on health searches has suggested that the two most important variable influencing the information consumers ultimately view and use are the processes by which they locate health information on the Internet and the evaluations they make regarding which web sites to review. [15] Hence it is of utmost importance to understand factors influencing online searches for information and to develop educational and technical approaches to maximize the quality and benefit of health searches.

Search engines have been identified as the preferred method of identifying information on the Internet but their use doesn't guarantee the retrieval of relevant websites. [1] While Google and HON rank as the best general and medical search engines respectively, a study assessing bone mineral density testing information from consumer health websites demonstrates that most of the relevant information is not accessible to the patients using the most popular Internet search engines. [1,5] Search results may vary due to the structure of content on health Web Sites, consumer search terminology and the use of paid placements by a search engine. [15]

The working of Google can be attributed to the PageRank that forms the basis of its software and a distributed network of thousands of low-cost computers that can carry out parallel processing and thereby speeding up the data processing. The three components of Google are: [21,22,23]

Googlebot: This is the web crawling robot that fetches and retrieves pages from the web and sends them to the Google indexer. This web crawler consists of many computers retrieving pages much faster than a web browser. By a technique called deep crawling this robot is capable of making massive and deep searches within the web. And it accomplishes this task by harvesting links from every page it encounters. To keep the

index current, Googlebot makes continuous recrawls at a rate proportional to the rate of change of web pages.

Google's Indexer: The indexer stores the text of web pages it receives from Googlebot and sorts them alphabetically by search term. Google doesn't index common words like the, is, on, or, of, how, why as well as some punctuation and multiple spaces to enhance search results. Also to improve performance all letters are converted into upper case.

Google's Query Processor: This includes the user interface, the query evaluator and the results formatter. Some of the factors it takes into account while determining the relevance to a query include the popularity of the page, the position and size of the search terms within the page and the proximity of search terms to one another on the page. All this has been termed as the PageRank for ranking of web pages by Google. It also applies machine-learning techniques like the spelling-correcting system to improve the performance of the search engine.

Google lends more priority to pages that have search terms near each other and in the same order as in the query. Using the advanced search options, Google allows users to restrict searches on the basis of where the query word appears. [21,22,23]

To address specifically to the needs of health information seekers there has been an upsurge in the number of health specific search engines. Healths on the Net Foundation along with the Molecular Imaging and Bioinformatics Laboratory at Geneva University Hospital have developed Multi-agent Retrieval Vagabond on Information Networks (MARVIN). This robot is the backbone behind the health-specific search engine MedHunt. [24]

MARVIN selectively indexes documents relevant to health and medical field unlike other search engines. Document relevance is computed considering the number of words from the medical dictionary that are present in the document as well as their associated weights. The weights have been assigned according to the relevance and specificity of the word to the medical field and reflect its local and global significance. [24]

Along with performing a full text indexing MARVIN stores additional information with each document such as: number of words, number of specific words, title, main text headers, most significant specific words, word position in the document, referenced documents and pages referencing the document, type, length, creation and last modification dates. When the user submits a query, MedHunt, the medical search engine, then uses this specific information. [24]

How well they work when compared to the general search engines in addressing consumers health information needs continues to be a matter of debate. According to a study on consumer information about androgen deficiency, there wasn't any significant difference between the quality of health information retrieved by medical search engines to that by general search engines. [1]

Although most studies assessing Internet use and quality of information have been performed focusing on specific health topics, there has been limited research comparing these engines from a more generic perspective of health care information consumers. We chose to address this topic from the point of view of consumers who haven't had any background training in medical literature and might be potential candidates to look up for information on the Internet regarding their own health or those of their family members. We also hypothesized that these are the topics that consumers may chose to investigate

with the anonymity of the Internet. We also planned to follow this with an independent evaluation of the quality of the web pages by physicians.

Materials and Methods:

The objective of this study was to evaluate and compare the search engines rather than the online health information seeker. For the purposes of this study we chose to compare general search engine, Google with two health-specific engines, Healix and MedHunt. Our goal was to determine whether search engines developed specifically for retrieving quality health information performed better than the currently available search engines both from the consumers' perspective as well as from the quality of content judged by health professionals. Six questions covering broad areas of health interest including Diabetes, Colon Cancer, Drug Prescriptions, Weight-loss, Alternative medicine and smoking were chosen for the purposes of this study. The topics were chosen considering the vast popularity they enjoyed amongst the current community of web based health information seekers. Multiple methods of data collection were employed including naturalistic observation of consumers using the search engines along with post-search questionnaires.

Participants in the study were mostly recruited at the site through posted flyers and included mostly healthy volunteers who had at least searched for health information once before. They had to be at least 21 years old and must not have had any background training or expertise in health care. There were a total of 66 participants recruited for this study on a first come first serve basis.

Table 1: Screening Questions

- Do you speak and read English at the 8th grade level (at least)? (Must answer YES)
- Are you comfortable using a computer to look for information on the Web? (Novice use is fine, but need to have used the Web before) (Must answer YES)
- Do you or a family member work for a company or organization involved in Internet search engines or online health information? (Must answer NO)

Naturalistic Observation of the Study Participants:

We conducted this study at the OHSU outpatient clinic from December 2004 to January 2005 with the main goal of establishing how consumers search the web, how successful they are, and what criteria they use to define their satisfaction when using a particular search engine. We had a demographically diverse sample of study subjects. There were twice as many females and males. The mean age was 36 years old. The sample was spread across ethnic categories, level of education and income (Table 2).

Table 2: Subjects' Demographics

Mean age 35.6 yrs \pm 12.5 yrs (range 21-70 yrs)

Gender: 66.7% female and 33.3% male

Ethnicity

- White, non-Hispanic 78.8%

- African American 10.6%
- Asian / Pacific Islander 7.6%
- Multiracial 1.5%
- Other 1.5%

Level of Education

- 8th grade or less 3.0%
- High school or GED 30.3%
- Up to 2 years of college 28.8%
- Up to Bachelors degree 25.8%
- Up to Masters degree 7.6%
- Professional or doctorate 4.5%

Income

- < \$20,000 31.8%
- \$20,000 -> \$50,000 40.9%
- \$50,000 -> \$75,000 15.2%
- \$75,000 -> \$100,000 9.1%
- \$100,000 + 3.0%

The study participants were also asked about their general health impression, whether they were responsible for arranging medical care for the household. In addition, they were asked about their previous experiences with computers and health-related searches, as well as which search engines they had previously used. These results are shown in Table 3.

Table 3: Search engines that the subjects had used before

– Google	93.9%
– Yahoo!	77.3%
– Ask Jeeves	31.8%
– AOL Search	21.2%
– HotBot	6.1%
– Vital Seek	1.5%
– Healia	0.0%

We gave six representative health questions to each of the 66 participants and asked them to use a particular search engine to find an answer to the best of their ability. These questions are listed in Table 4. The ordering of the questions and the search engines were pseudo-randomized, so that each participant used each of the three search engines twice while answering a total of all six questions. The participants were instructed to search as they would at home, if they had a similar question of interest and to stop when they felt comfortable with a sufficiently good answer. We gave them a time limit of 10 minutes per question. Participants were encouraged to think aloud while performing the search

and use the same techniques they would use while searching the Internet from their home.

Table 4: Health Questions Used

1. Suppose your spouse is suffering from mild depression. Your friend recently told you that St. John's Wort is helpful for depression. Is there any evidence to support the claim that St. John's Wort is effective for treating mild depression?
2. Your doctor just prescribed cipro (ciprofloxacin) pills for your throat infection. Are there any foods you should avoid when taking this drug?
3. Your three-year-old child has just burned himself by accidentally touching the hot oven. Can you find out how to provide first aid for a minor burn?
4. You have decided to cut down on your Fat intake to improve your health. How many grams of fat are there in a McDonalds' Big Mac sandwich?
5. Your father died at the age of 60 of colon cancer. Is it true that people with a family history of colon cancer are more likely to get colon cancer than those without a family history of it?
6. Your friend is a cigarette smoker and has decided to quit smoking. Can you find a guide about quitting smoking for her to use? (She dropped out of high school and reads at about the 8th grade level).

The study was performed with two computers using special software to record the pages visited, search terms used, the order of pages retrieved, pages selected by the subject and number of pages visited until the subject found a satisfactory answer. In addition two researchers took field notes about whether subjects used advanced search features, whether they used single/multiple sites to obtain an answer and whether they had any confusion or difficulties.

Post-Search Questionnaires:

We asked the participants to answer a standard questionnaire each time they found a satisfactory answer to the search question or they were timed-out by the researcher. These questions included their general impression of the usability of the search engine, the ease of use, and the degree of relevancy of the search engine's results to that particular question and how satisfied they felt with the results retrieved by that engine (see Table 5). The responses were measured on a five point Likert scale.

Table 5: Post Search Questionnaire

What was your general impression of the usefulness of this search engine for finding health information?

Was this search engine easy to use?

How relevant were the results of your search using this Web site in answering the

question?

How satisfied were you with the results returned by the search engine?

We followed the set of six searches with a final debriefing session where each individual participant was asked to answer certain open-ended questions, such as

- *What is important to you when you use a search engine to get for health information?*
- *Do you have any general comments or suggestions for the system developers?*
- *Which search engine did you prefer and why? (Score 1-10)*

Evaluation by Physicians:

In addition to testing the performance of the search engines from the health care consumer's point of view, we also tested the quality of the returned results, as judged by health care professionals. We analyzed the search terms used by the patient participants from the log record of the recording software. We chose the top 2 keyword phrases used by participants in answering each question and reproduced the results to be rated by three physicians independently. We collected the top 10 Web sites for each of the two keyword phrases for each search engine and all 6 questions. This included about 60 Web pages for each question. The physicians were asked to give a score to each Web page ranging from 1 being the poorest to 10 being excellent on each of the following parameters: relevancy of the content in answering the specific question, accuracy of the

information on the particular web page and trustworthiness of the source of information. They were asked to rate the linked page only and not to pursue several links to find an answer to the health question.

Results and Analysis:

The results of the study were analyzed by performing paired sample test using SPSS software. The *first phase* involved analysis of data gathered from 66 consumers as study participants. After each attempt to find an answer to one of the six health questions using one of the three search engines (pseudo-random presentation), the participants were asked to rate the usefulness of the search engine, its ease-of-use, the relevance of the material returned, and their overall satisfaction with the experience. We used a Likert scale from 0-4, with 4 being the best, to measure their responses. In general, from the consumer participant's perspective, the Google search engine was the most preferred, followed by Healia and then MedHunt. The scores for each dimension are shown below in Figure 1. For our analysis of the significance of the results, we used a 2-tailed paired t-test to compare the performance of each pair of search engines on each of the dimensions. The significance results are shown in Table 6. Both Google and Healia were always significantly better than MedHunt ($p < .005$). However, Google was perceived to be significantly better than Healia on only one dimension – usefulness ($p = .033$). We found a tendency toward a preference for Google with respect to ease-of-use ($p = .062$), but the scores were similar for the relevance of the results and overall satisfaction.

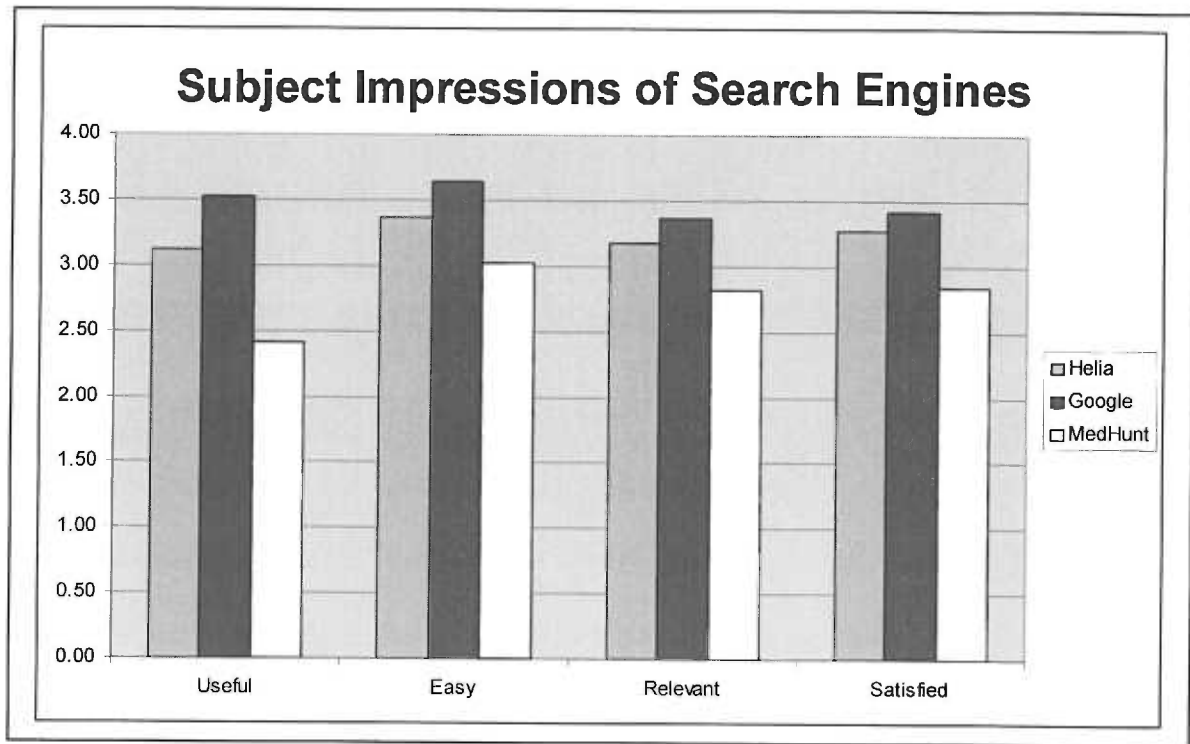


Figure 1: Consumer participant feedback on their experiences using the three search engines: Healia, Google, and MedHunt. Responses were measured on a scale of 0 to 4, with a score of 4 being best.

Table 6: Statistical results for a paired t-test analysis of each combination of search engines, comparing results on usefulness, ease-of-use, relevance of results and overall satisfaction.

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Google Usefulness - Healia Usefulness	.3205	.90675	.14520	.0266	.6144	2.207	38	.033
Pair 2	Healia Usefulness - MedHunt Usefulness	.8333	.87610	.14029	.5493	1.1173	5.940	38	.000
Pair 3	Google Usefulness - MedHunt Usefulness	1.1538	.99442	.15923	.8315	1.4762	7.246	38	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Google Ease of Use - Healia Ease of Use	.1923	.62424	.09996	-.0100	.3947	1.924	38	.062
Pair 2	Healia Ease of Use - MedHunt Ease of Use	.4744	.57295	.09175	.2886	.6601	5.170	38	.000
Pair 3	Google Ease of Use - MedHunt Ease of Use	.6667	.73747	.11809	.4276	.9057	5.645	38	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Google Relevance - Healia Relevance	.1538	.74475	.11926	-.0876	.3953	1.290	38	.205
Pair 2	Healia Relevance - MedHunt Relevance	.4872	.73884	.11831	.2477	.7267	4.118	38	.000
Pair 3	Google Relevance - MedHunt Relevance	.6410	.75174	.12037	.3973	.8847	5.325	38	.000

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Google Satisfaction - Healix Satisfaction	.1282	.84072	.13462	-.1443	.4007	.952	38	.347
Pair 2	Healix Satisfaction - MedHunt Satisfaction	.5385	.83022	.13294	.2693	.8076	4.050	38	.000
Pair 3	Google Satisfaction - MedHunt Satisfaction	.6667	.86855	.13908	.3851	.9482	4.793	38	.000

Throughout the experiment, while the consumer participants were using the various search engines, the researchers kept qualitative notes, noting problems, barriers and areas of confusion. The researchers specifically documented the possible use of advanced search features (i.e., ctrl-F within the Web browser). Several participants used the search function within a resultant web page if this function was available. One area of confusion was that many participants looked for the answer to the question in the results summary page instead of clicking on links to resultant url's to find the answer. We found that a common comment from the participants was that they liked the Google search engine because they were already used to it. A summary of our qualitative finding is shown in Table 7 below.

Table 7: Summary of Qualitative Findings for Each Search Engine

Healia

- Many participants used the advanced search features in Healia
- Many participants felt that Healia was better suited for professionals as opposed to lay people

Google:

- At least a couple of the participants clicked on Google's "news" link since it appeared to be the first result
- Several users clicked on a "sponsored link" but didn't find the answer to the question on these sites
- Participants almost never used the advanced search features in Google.

MedHunt

- Participants commented on MedHunt far more often; most comments were negative
- Many participants used the advanced search features in MedHunt
- Many participants felt that the MedHunt interface was confusing
 - Clicking on HON code link thinking that it was a result
 - Not knowing which portion of a result is the link to the website
- Many participants felt that MedHunt was better suited for professionals, as opposed to lay people

Whenever participants used quotes in their search strategy, MedHunt returned zero results

At the end of consumer participants' study session, we asked them to give an overall rating to their experiences with each search engine on a scale of 1 to 10 (10 being best). Figure 2 shows the participants' scores for each of the search engines. Overall, from the consumer perspective, Google was the preferred search engine, followed by Healia and last by MedHunt. Again, we used a 2-tailed paired t-test to look for significant differences between pairs of search engines. Table 8 shows that Google was perceived to be significantly better than both Healia ($p < .005$) and MedHunt ($p < .005$). Healia was also perceived to be significantly better than MedHunt ($p < .01$).

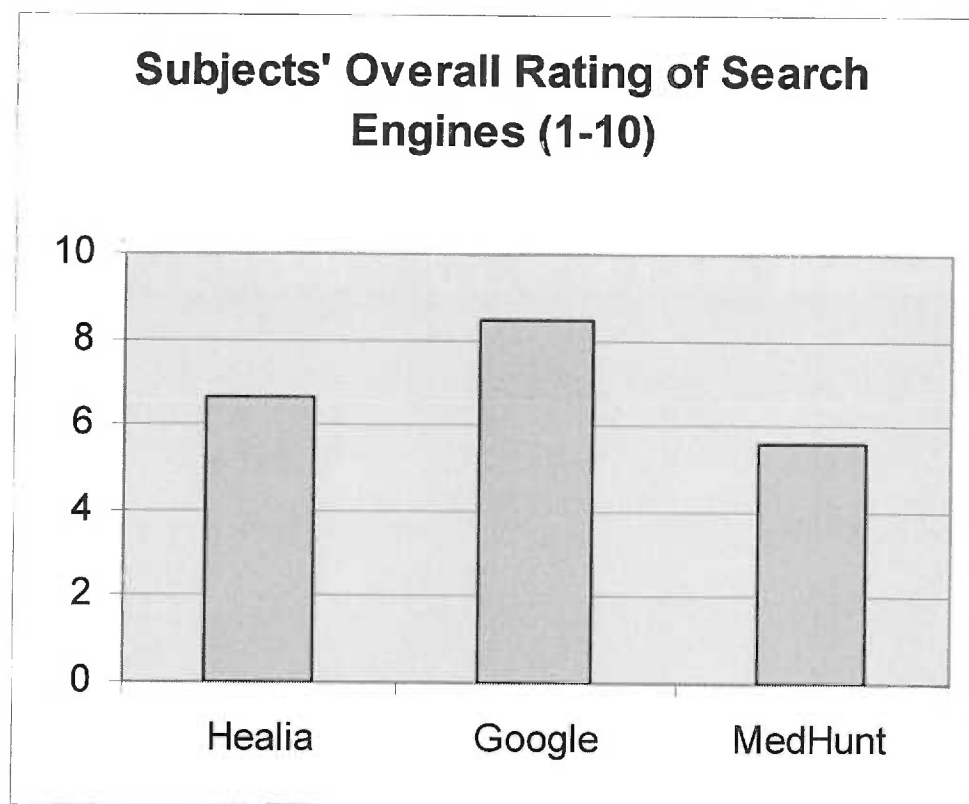


Figure 2: Consumer subjects' ratings of the search engines on a scale of 1 to 10 (10 best)

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Google Score - Helia Score	1.80	2.934	.367	1.06	2.53	4.899	63	.000
Pair 2	Helia Score - MedHunt Score	1.05	3.036	.380	.29	1.81	2.758	63	.008
Pair 3	Google Score - MedHunt Score	2.84	3.252	.407	2.03	3.66	6.995	63	.000

Table 8: Analysis of significant differences in overall ratings between search engines.

We also performed a subgroup analysis to see if any of the search engines were more helpful to participants based on demographic differences. We looked for possible significant differences in scores based on gender, age, income, ethnicity and education. In this analysis we found that the mean scores given to Helia by participants who were 8th graders were significantly lower than scores by those who had a Master's level of education (See Appendix D).

Phase 2 Evaluation: Search Engine Quality as Measured by Clinicians

In the second phase of our evaluation, we asked three clinicians to evaluate the quality of the Web sites returned by the three search engines during Web searching by the consumer participants. To do this, we analyzed all of the keyword phrases used by the consumer participants for each of the six health questions and selected the two most commonly used search terms for each. We then entered those keyword phrases into each of the three

search engines and selected the top ten urls returned for each. After randomizing the presentation of these urls, we asked the health professionals (three physicians with an interest in consumer health informatics) to evaluate the quality of the Web sites. Figure 3 shows the instructions given to the clinicians for a sample question, along the beginning of the file for entering quality assessments for each Web site. In a separate record, we maintained a field marking which search engine had returned that site and the order in which it was presented to the consumer. The clinicians rated the relevancy of the material to the question the consumers had seen, the accuracy of the information, and the trustworthiness of the information on a scale of 1 to 10 (with 1 being very poor and 10 being excellent).

Figure 3: Portion of the assessment screen the clinicians used to evaluate the quality of the Web sites, scoring relevancy, accuracy and trustworthiness on a scale of 1 (very poor) to 10 (excellent).

On a scale of 1 (very poor) to 10 (excellent) with 5 being an average score, rate each page according to

Relevancy - Did it answer the question? Was it on topic?

Accuracy - Is the material on the page accurate? Is it a correct answer to the question?

Trustworthiness - Do you trust the source of the information? Confidence in stated process for review and updates.

Please rate the linked page only.
However, you may look at other pages on the site (e.g., "about us" page) if it will help your assessment of the reliability and trustworthiness of the source.

Question: Your friend is a cigarette smoker and has decided to quit smoking.
Can you find a guide about quitting smoking for her to use?
(She dropped out of high school and reads at about the 8th grade level).

Relevancy	Accuracy	Trustworthiness	COMMENTS (optional)	URL
7	7	7		http://my.webmd.com/content/article/1/1700_50723.htm
7	7	7		http://my.webmd.com/content/article/51/40657.htm
8	7	7		http://my.webmd.com/content/healthwise/68/17000
8	6	6		http://quitsmoking.about.com/
8	6	6		http://quitsmoking.about.com/

In addition to the six health questions presented to the consumer participants, we added results for typical keyword phrases to three additional health questions, designed to test additional features of the Healia search engine. The additional health questions were:

- Do any herbal supplements help cure cancer?*
- How does Viagra compare to Levitra?*
- Which types of diet pills are effective for weight loss?*

In our analysis of the quality ratings by the health professionals, we first looked for consistency among raters. Figures 4-6 show that each health professional rated the Healia search engine more highly than Google and MedHunt. Although these individual differences weren't statistically significant, the ordering of the ratings remained consistent across raters for each dimension (accuracy, relevance, and trustworthiness).

Figure 4: Accuracy ratings by each health professional for the three search engines.

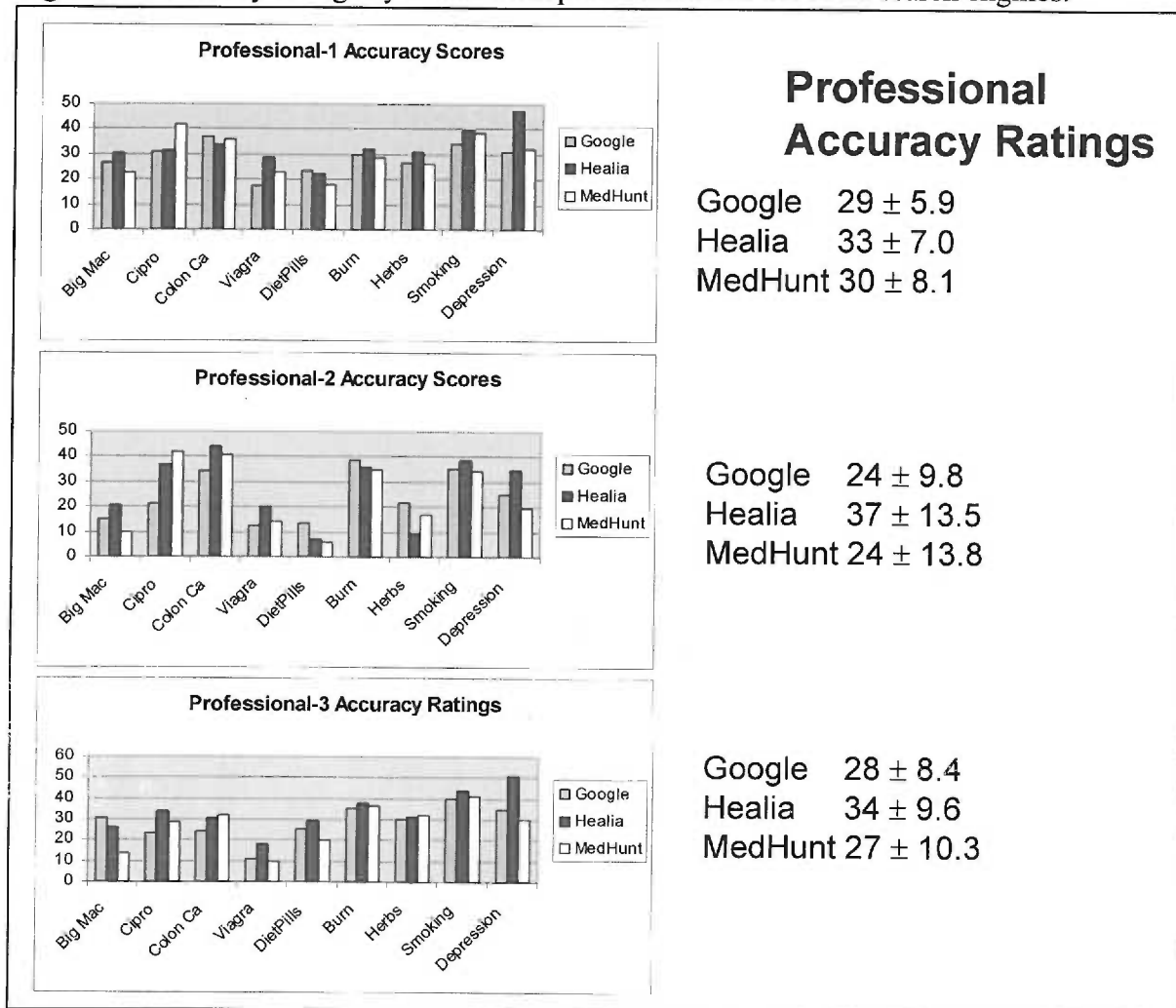


Figure 5: Relevancy ratings by each health professional for the three search engines.

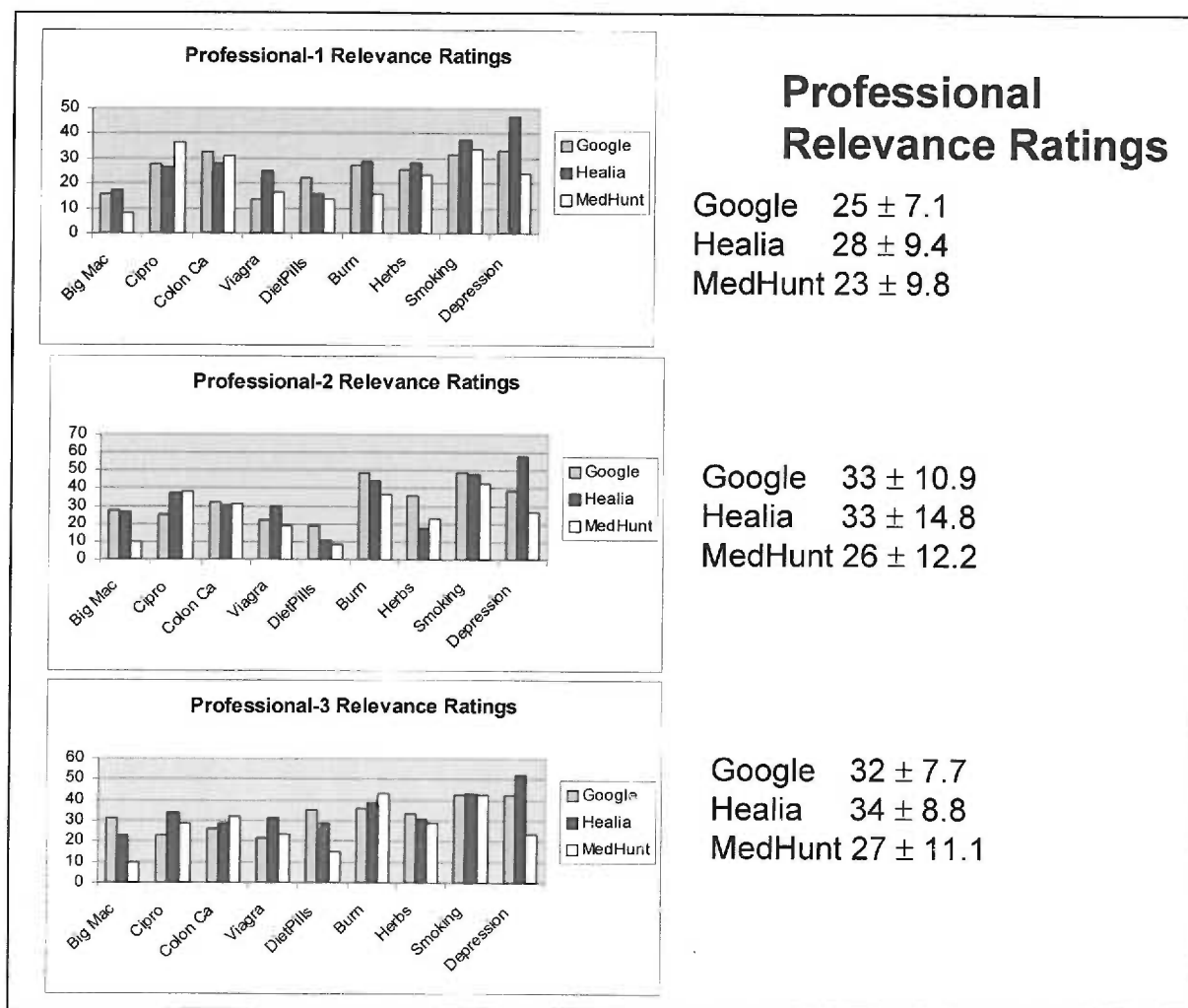
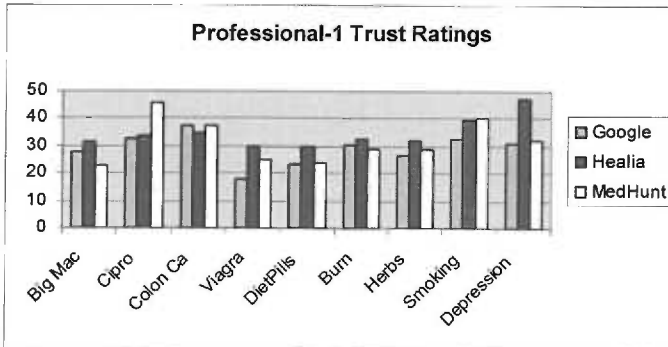
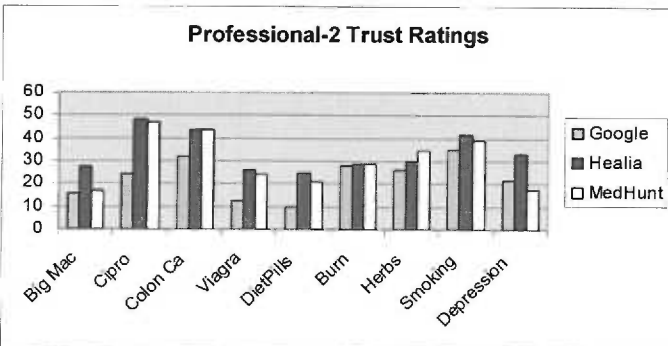


Figure 6: Trustworthiness ratings by each health professional for the three search engines.

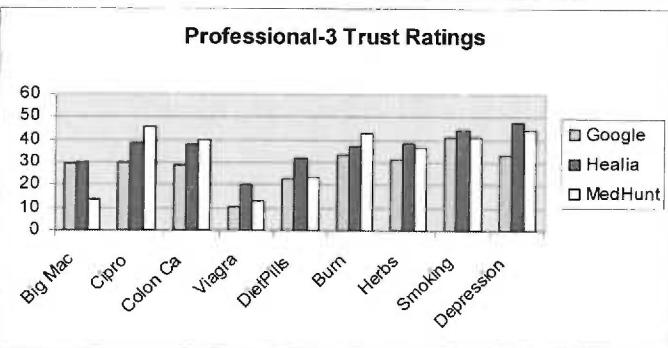


Professional Trust Ratings

Google 29 ± 5.7
 Healia 35 ± 5.6
 MedHunt 32 ± 8.0



Google 23 ± 8.6
 Healia 34 ± 8.7
 MedHunt 30 ± 11.4



Google 29 ± 8.4
 Healia 36 ± 8.0
 MedHunt 33 ± 13.0

Because of the consistency in the ordering of the results across health professional raters as shown in Figures 4-6, we were able to average scores across raters for further analysis. We also averaged across keyword phrases and health questions.

Figure 7: Averaged health professional ratings of accuracy, relevance and trustworthiness of Web sites returned by each of the search engines.

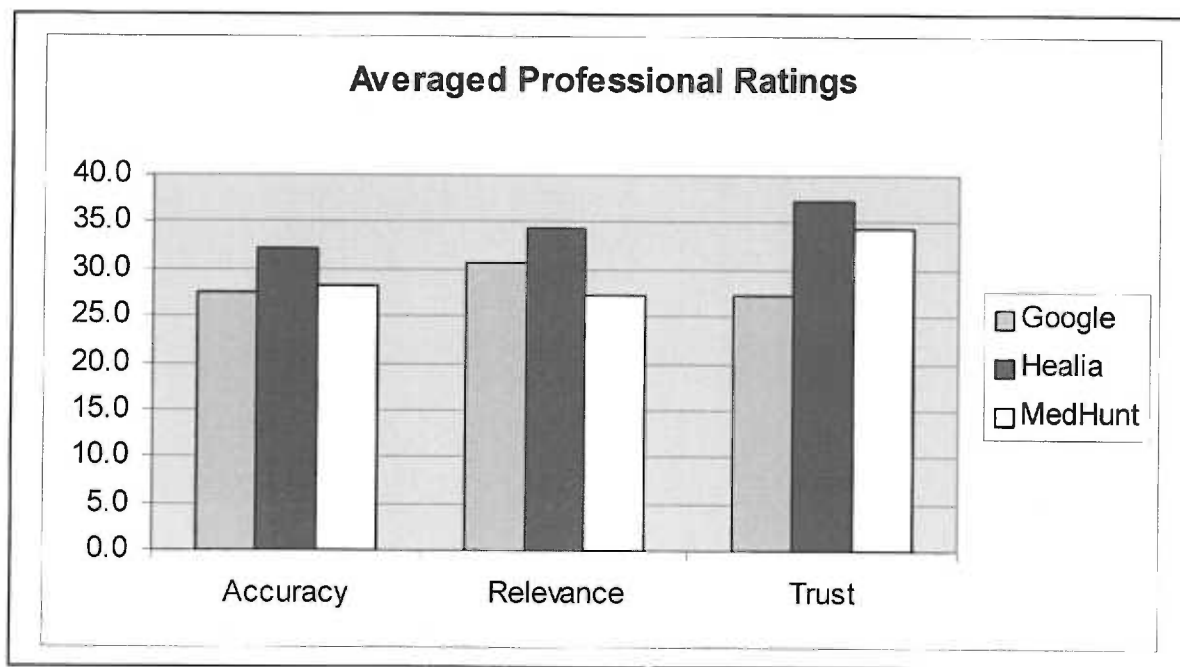
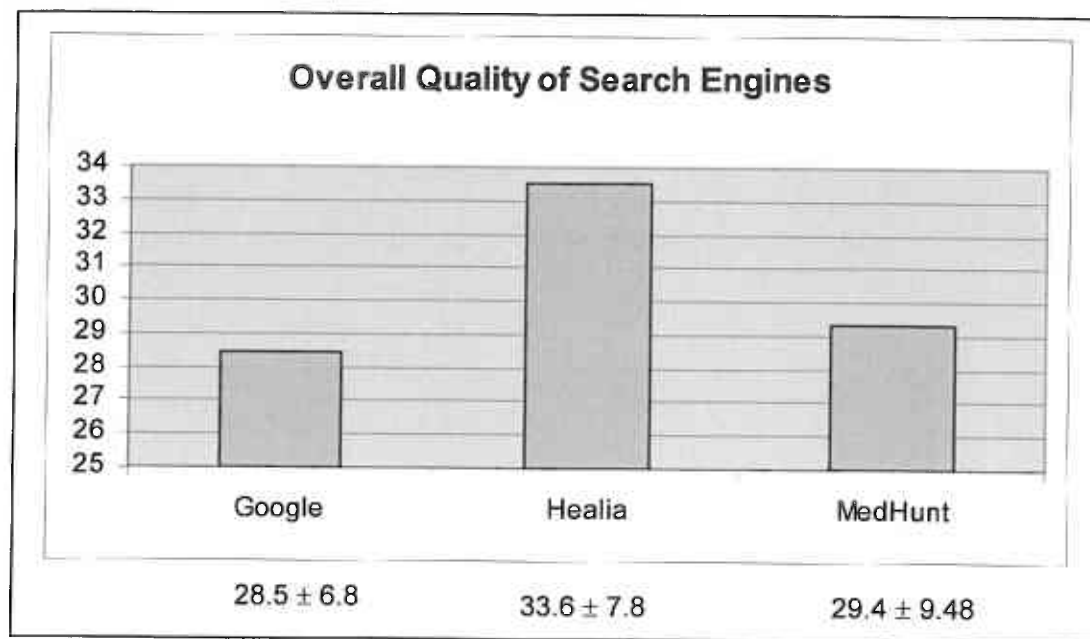


Figure 7 shows our interim analysis with the professionals' ratings for accuracy, relevance, and trustworthiness separately. Healia is rated more highly than the other search engines on each dimension, although the difference is not statistically significant with our number of subjects.

Figure 8: Overall quality judgments from health professionals



For the analysis shown in Figure 8, we combined the three quality dimensions (equal weighting). Again, although the differences in quality between search engines is not statistically significant at the $p < .05$ level, there is a clear tendency for Healia to perform better than the other search engines.

Summary and Conclusions

The results of this study indicate that consumers search for and interpret health-related information in a different way than is implicitly assumed by investigators trying to assess quality of health information on the web specific to a particular topic of interest. In general, participants were very successful in finding answers to health questions by trying various search terms, exploring the first few hits by cursory examination of the content of the page, and iteratively refining their search strategy. This study is, to our knowledge,

one of the very few studies to investigate the retrieval strategies of common people searching for health information on the web as well as to measure their attributes of satisfaction and usefulness.

The study has several important limitations. Firstly, we observed only a small number of participants in a somewhat artificial environment under experimental conditions. We cannot rule out the possibility that participants may have felt time pressure and thus put less emphasis on checking the quality and source. We tried to minimize this by explicitly instructing them to take their time and try to retrieve the correct answer rather than a quick answer. Secondly, as answering the questions did not have any direct impact on the participants, people in a real setting with a greater stake in the outcome of the search might care more about quality and therefore more actively look for markers of quality. Thirdly, we do not know to what degree the sample we tested was representative, but it is likely that we had a relatively experienced sample that was well spread across the age groups, ethnicity and income. We chose to exclude health professionals from participating in this study since we were more interested in observing the lay consumers who weren't very familiar with medical vocabulary and search terminology.

We found that Healia was preferred to the existing competing health search engine by consumers on dimensions of usefulness, ease-of-use, relevance and overall satisfaction. However, the highly successful general search engine, Google, was preferred to both Healia and MedHunt on the dimension of usefulness. In addition, there was a significant ordering of results when we asked for a final rating of search engines at the end of the experiment, with consumers preferring Google to Healia, and Healia to MedHunt. From our qualitative findings, it seems highly likely that a major factor was that our consumer

participants had previous experience with Google. In addition, the Google search engine has a very simple interface.

Our results from a variety of analyses of the clinicians' ratings of quality of the Web sites that were presented to the consumers by the three search engines showed a general trend that on the dimensions of accuracy, relevance and trustworthiness, Healia was consistently rated more highly than the other search engines. With a larger number of raters we may have been able to show statistical significance of this finding.

More observational studies are required to further understand the behavior of health information consumers and to design and evaluate educational and technological innovations for guiding consumers to high quality health information on the web. This would be the definite challenge in this rapidly advancing era of "Health Consumer Informatics".

References:

- 1) D.Ilic, T.L.Bessel, C.A.Silagy and S.Green. Specialized medical search-engines are no better than general search-engines in sourcing consumer information about androgen deficiency, *Human Reproduction* Vol.18, No.3 pp. 557±561, 2003
- 2) Anna Gagliardi and Alejandro R Jadad. Examination of instruments used to rate quality of health information on the Internet: chronicle of a voyage with an unclear destination, *BMJ* 2002; 324; 569-573
- 3) Gunther Eysenbach and Christian Kohler 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. *BMJ* 2002; Volume 324
- 4) Matthias Wjst When Air Is Rare: Behind the Scenes Of an Asthma Web Site *Journal of Asthma*, 38(5), 399–404 (2001)
- 5) Carolyn J. Green, Armin'ee Kazanjia and Diane Helmer. Informing, advising, or persuading? An assessment of bone mineral density testing information from consumer health websites *International Journal of Technology Assessment in Health Care*, 20:2 (2004), 156–166
- 6) Funda Meric, Elmer V Bernstam, Nadeem Q Mirza, Kelly K Hunt, Frederick C Ames, Merrick I Ross, Henry M Kuerer, Raphael E Pollock, Mark A Musen and Eva Singletary Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites *BMJ* 2002; 324; 577-581

- 7) Judith L Bader MD; Mary Frances Theofanos , MS Searching for Cancer Information on the Internet: Analyzing Natural Language Search Queries Journal of Medical Internet Research Volume 4, Issue 4, Article e31
- 8) MATTHEW I. KIM AND PAUL W. LADENSON Diabetes on the Internet: Resources for Patients The Journal of Clinical Endocrinology & Metabolism 87(8): 3523–3526 (2002)
- 9) Katherine E Waack, Michael E Ernst, and Mark A Graber Informational Content of Official Pharmaceutical Industry Web Sites about Treatments for Erectile Dysfunction The Annals of Pharmacotherapy _ 2004 December, Volume 38
- 10) E.C.Haagen, W.Tuil, J.Hendriks, R.P.J.de Bruijn, D.D.M.Braat and J.A.M.Kremer.Current Internet use and preferences of IVF and ICSI Patients Human Reproduction Vol.18, No.10 pp. 2073±2078, 2003
- 11) Anthony G. Crocco, Miguel Villasis-Keever and Alejandro R. Jadad.Analysis of Cases of Harm Associated With Use of Health Information on the Internet (Reprinted) JAMA, June 5,2002-Vol 287, No. 21
- 12) Gretchen K. Berland, Marc N. Elliott, Leo S. Morales, Jeffrey I. Algazy, Richard L. Kravitz, Michael S. Broder, David E. Kanouse, Jorge A. Muñoz, Juan-Antonio Puyol Marielena Lara, Katherine E. Watkins, Hannah Yang and Elizabeth A. McGlynn. Health Information on the Internet Accessibility, Quality, and Readability in English and Spanish.JAMA, May 23/30, 2001—Vol 285, No. 20 (Reprinted)
- 13) Gunther Eysenbach, The Impact of the Internet on Cancer Outcomes CA Cancer J Clin 2003; 53:356–371

- 14) Don Fallis and Martin Fricke. Indicators of Accuracy of Consumer Health Information on the Internet: A Study of Indicators Relating to Information for Managing Fever in Children in the Home. Journal of the American Medical Informatics Association Volume 9 Number 1 Jan / Feb 2002
- 15) Liza Greenberg, Guy D'Andrea and Dan Lorence. Setting the Public Agenda for Online Health Search A White Paper and Action Agenda URAC December, 2003
- 16) Gunther Eysenbach, John Powell, Oliver Kuss and Eun-Ryoung. Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web:A Systematic Review. (Reprinted) JAMA, May 22/29, 2002-Vol 287, No.20
- 17) Mike Benigeri and Pierre Pluye. Shortcomings of health information on the Internet. Health Promotion International, Vol. 18, No. 4, 381-386, December 2003
- 18) Suzanne Dickerson, Amber M. Reinhart, Thomas Hugh Feeley, Rakesh Bidani, Elles Rich, Vinod K. Garg, Charles O. Hershey. Patient Internet Use for Health Information at Three Urban Primary Care Clinics Journal of the American Medical Informatics Association Volume 11 Number 6 Nov / Dec 2004
- 19) Kristina Eriksson-Backa. Who uses the Web as a health information source? Health Informatics Journal 2003 Vol 9(2): 93-101
- 20) Jean Harvey-Berino, Stephen Pintauro, Paul Buzzell, and Elizabeth Casey Gold. Effect of Internet Support on the Long-Term Maintenance of Weight Loss. OBESITY RESEARCH Vol. 12 No. 2 February 2004
- 21) http://www.googleguide.com/google_works.html, last visited 23rd March, 2005

- 22) www.google.com/technology, last visited 23rd March, 2005
- 23) <http://www.google.com/technology/pigeonrank.html>, last visited 23rd March, 2005
- 24) C. Boyer, O. Baujard, V. Baujard, S. Aurel, M. Selby and R.D Appel. Health on the Net automated database of health and medical information. International Journal of Medical Informatics 47 (1997) 27-29

Appendix A: Study Recruitment Flyer

You are invited to take part in a research study on how to help patients search for health information on the Web.

You are eligible for this study if you are at least 21 years old and have used a computer to search the Web at least once in the past year.

What We'll Ask You to Do:

- **We will first ask you to fill out a short background survey.**
- **You will then be given 6 health questions, and asked to find answers to these questions using 3 different Web search tools.**
- **We will ask for your feedback on your experience using each of the Web search tools, and at the end, we will ask you to compare them.**

♦ **The study will take about 45 minutes.**

♦ **You would be paid \$25 for completing the study.**

**To Sign Up: Contact the researcher at the table with
computers.**

We will be happy to answer any questions you have.

Department of Medical Informatics and Clinical Epidemiology

Oregon Health & Science University

3181 SW Sam Jackson Park Road

Portland, Oregon 97239

IRB # _____

Appendix B: Study Consent Form

IRB# e565

Protocol Approval Date: 11/16/2004

OREGON HEALTH & SCIENCE UNIVERSITY

Consent Form

TITLE: Health Search Tool Evaluation

PRINCIPAL INVESTIGATOR: Holly B. Jimison, PhD (503) 418-2277

SPONSOR: Evalumetrix, Inc. and the National Cancer Institute

PURPOSE:

The purpose of this study is to evaluate the performance of a newly developed computer tool for helping consumers search for health information on the Web, as compared to commercially available Web searching tools.

You have been invited to be in this research study because you are an adult who is familiar with searching for information on the Web.

We plan to study about 70 subjects to help us compare the performance of the 3 Web search tools.

PROCEDURES:

You can do this study today. It will take about 45 minutes.

We will first ask you to fill out a short survey on some background information, such as age, gender, ethnicity, general health experience, and familiarity with computers.

This will take 5 or 10 minutes.

We will then provide you with 6 health questions to try and answer using the Web.

The research assistant will set the computer to start with a different search tool for each question. After you feel that you have found an answer to each health question, we will ask you about your experience using the Web search tool.

After you have tried all 3 Web search tools, we will ask you to compare them.

RISKS AND DISCOMFORTS:

We do not expect this study to harm you. You may refuse to answer any of the questions we ask.

BENEFITS:

You will not personally benefit from being in this study. However, by serving as a subject, you may help us learn how to help other patients search for health information in the future.

ALTERNATIVES:

You may choose not to be in this study.

CONFIDENTIALITY:

Forms with your name (consent form and payment form) are kept separate from other forms. We will use an ID number instead of your name on the survey and notes. Our records of your answers to our questions will not be linked to you. All data will be kept in a locked and secure location. Research records may be reviewed and copied

by the sponsors, the OHSU Institutional Review Board, and the Office for Human Research Protections.

We will not use your name or your identity for publication or publicity purposes.

COSTS:

We will pay you \$25 for completing the study.

LIABILITY:

It is not the policy of the U.S. Department of Health and Human Services, or any federal agency funding the research project in which you are participating to compensate or provide medical treatment for human subjects in the event the research results in physical injury.

The Oregon Health & Science University is subject to the Oregon Tort Claims Act (ORS 30.260 through 30.300). If you suffer any injury and damage from this research project through the fault of the University, its officers or employees, you have the right to bring legal action against the University to recover the damage done to you subject to the limitations and conditions of the Oregon Tort Claims Act. You have not waived your legal rights by signing this form. For clarification on this

subject, or if you have further questions, please call the OHSU Research Integrity Office at (503) 494-7887.

PARTICIPATION:

Holly Jimison (503) 418-2277 has offered to answer any questions you may have about this study. If you have any questions regarding your rights as a research subject, you may contact the OHSU Research Integrity Office at (503) 494-7887.

You do not have to join this or any research study. If you do join, and later change your mind, you may quit at any time. If you refuse to join or withdraw early from the study, there will be no penalty or loss of any benefits to which you are otherwise entitled.

You will be given a copy of your signed consent form.

SIGNATURE:

Your signature below indicates that you have read this entire form and that you agree to be in this study.

Printed Name of Subject

Date

Signature of Subject

Signature of Person Obtaining Consent

Date

Appendix C: Sample Screen Shots of Study Assessment Tool (Consumer Participants)

Question

Your three-year-old child has just burned himself by accidentally touching the hot oven.
Can you find out how to provide first aid for a minor burn?

**To go to the search engine for this question,
please click on the link below**

[Healia](#)

Question

Your father died at the age of 60 of colon cancer.
Is it true that people with a family history of colon cancer are more likely to get colon cancer than those without a family history of it?

**To go to the search engine for this question,
please click on the link below**

[MedHunt](#)

Table 1: Sub-analysis of differences in scores for the three Search Engines according to Gender (Significance is 2-tailed)

	Sex	N	Mean Score	Std. Deviation	Significance (2-tailed)
Healia Score	Male	22	6.59	1.968	0.902
	Female	42	6.67	2.486	
Google Score	Male	22	8.36	1.432	0.807
	Female	42	8.48	1.877	
MedHunt Score	Male	22	5.73	2.313	0.769
	Female	42	5.52	2.770	

Table 2: Sub-analysis of differences in scores for the three Search Engines according to Age (Significance is 2-tailed)

		N	Mean Score	Std. Deviation	Significance (2-tailed)
Healia Score	20's	26	6.88	2.179	
	30's	12	5.83	2.758	
	40's	12	6.67	1.969	
	50's	9	6.67	2.236	
	60's	2	8.50	2.121	
	70's	1	8.00		

Google Score	Total	62	6.68	2.245	0.627
	20's	26	8.65	1.441	
	30's	12	8.83	1.467	
	40's	12	8.25	2.454	
	50's	9	7.56	1.590	
	60's	2	9.00	.000	
	70's	1	9.00	.	
	Total	62	8.47	1.686	
MedHunt Score					0.562
	20's	26	6.27	2.585	
	30's	12	4.00	2.216	
	40's	12	6.00	2.486	
	50's	9	5.67	2.646	
	60's	2	5.00	4.243	
	70's	1	8.00	.	
	Total	62	5.68	2.604	
					0.188

Table 3: Sub-analysis of differences in scores for the three Search Engines according to Level of Education (Significance is 2-tailed)

		N	Mean Score	Std. Deviation	Significance
Healia Score	8 th Grade or less	2	3.00	2.828	0.014
	High school or GED	20	6.45	2.724	
	Associates	19	7.32	1.827	
	Bachelors	15	6.13	1.846	
	Masters	5	8.80	1.304	
	Professional/Doctorate	3	5.00	.000	
	Total	64	6.64	2.305	
Google Score	8 th Grade or less	2	9.00	1.414	0.795
	High school or GED	20	8.60	1.903	
	Associates	19	8.00	1.915	
	Bachelors	15	8.80	1.014	
	Masters	5	8.40	1.949	
	Professional/Doctorate	3	8.00	2.646	
	Total	64	8.44	1.726	
MedHunt Score	8 th Grade or less	2	2.50	2.121	
	High school or GED	20	6.30	2.408	
	Associates	19	5.63	3.041	
	Bachelors	15	5.93	1.981	
	Masters	5	4.40	2.510	
	Professional/Doctorate	3	3.00	2.000	

Total				
	64	5.59	2.605	
				0.122

Table 4: Sub-analysis of differences in scores for the three Search Engines according to Ethnicity (Significance is 2-tailed)

		N	Mean Score	Std. Deviation	Significance
Healia Score	White, non-Hisp	50	6.76	2.317	0.784
	African-American	7	5.86	2.911	
	Asian/Pacific Islander	5	6.60	1.673	
	Multiracial	1	8.00	.	
	Other	1	5.00	.	
	Total				
		64	6.64	2.305	
Google Score	White, non-Hisp	50	8.60	1.690	
	African-American	7	8.14	1.864	
	Asian/Pacific Islander	5	7.20	2.049	
	Multiracial	1	8.00	.	
	Other	1	9.00	.	

MedHunt Score	Total	64	8.44	1.726	0.507
	White, non-Hisp	50	5.70	2.652	
	African-American	7	6.00	2.708	
	Asian/Pacific Islander	5	5.00	2.345	
	Multiracial	1	3.00		0.658
	Other	1	3.00		
	Total	64	5.59	2.605	

Table 5: Sub-analysis of differences in scores for the three Search Engines according to Income (Significance is 2-tailed)

		N	Mean	Std. Deviation	Significance
Healia Score	Less than 20,000	21	6.48	2.695	
	20 to 50,000	27	6.30	2.053	
	50 to 75,000	8	7.38	1.598	
	75 to 100,000	6	8.67	1.633	
	Over 100,000	2	4.00	1.414	

Google Score	Total	64	6.64	2.305	0.061
	Less than 20,000	21	8.24	1.640	
	20 to 50,000	27	8.44	2.006	
	50 to 75,000	8	8.88	1.246	
	75 to 100,000	6	8.50	1.761	
	Over 100,000	2	8.50	.707	
	Total	64	8.44	1.726	
MedHunt Score	Less than 20,000	21	5.29	2.610	0.942
	20 to 50,000	27	5.93	2.541	
	50 to 75,000	8	6.63	2.560	
	75 to 100,000	6	4.33	3.077	
	Over 100,000	2	4.00	1.414	
	Total	64	5.59	2.605	
					0.396

