Proceed with Caution:
A Report on the Quality of Health Information on the Internet

Complete Study

May 2001

Prepared for the
California HealthCare Foundation

by
RAND Health

Proceed with Caution: A Report on the Quality of Health Information on the Internet is a study in three parts: Report Summary, Consumer Report, and Complete Study. To receive the other reports or additional copies of this report, contact the California HealthCare Foundation's publications line at (510) 587-3199 or visit us online (http://ehealth.chcf.org).

RAND is a nonprofit institution that helps improve policy and decision making through research and analysis. RAND serves the public interest by applying its collective expertise to illuminate the most challenging societal problems, to identify effective courses of action, and to devise innovative solutions.

The California HealthCare Foundation (CHCF), located in Oakland, is an independent health care philanthropy committed to making the health care system work better for the people of California. Created in 1996, CHCF focuses on critical issues in the following areas: health policy and regulation, health care delivery and financing, insurance and the uninsured, Medicare, Medi-Cal, and Healthy Families, quality of care, and ehealth. Grantmaking is directed at innovative research, development of model programs, and projects that will yield meaningful policy recommendations.

ISBN 1-929008-63-5

CALIFORNIA HEALTHCARE FOUNDATION
476 Ninth Street
Oakland, CA 94607
tel: (510) 238-1040
fax: (510) 238-1388
www.chcf.org

Acknowledgments

This study was supported by grants from the California HealthCare Foundation (Berland, PI) and the Robert Wood Johnson Foundation Clinical Scholars Program (Berland and Algazy). A shorter version of this study is also published in the May 23, 2001, issue of the *Journal of the American Medical Association* (JAMA, May 23/30, 2001-Vol. 285, No. 20).

We appreciate Dan Durazo's advice in selecting Spanish search engines and Web sites. Throughout the project we have benefited from the support and guidance of Sam Karp, Chief Information Officer and Ann F. Monroe, Director, Quality Initiative, both of the California HealthCare Foundation.

The third chapter would not have been possible without the diligent efforts of the 34 physicians who reviewed the material we selected from the Web sites being evaluated. They include: Evan Alley, MD, PhD; David Axelrod, MD, MBA; Patricia Barretto, MD; Deborati Borschel, MD; Marcia Brose, MD, PhD; Marcia Chantler, MD; Peter Cram, MD; Rebecca Elstrom, MD; Penelope Ewbank, MD; Renee Gerrero, MD; Katherine Goodchild, MD; Jini Han, MD; Michele Heisler, MD; David Lanfear, MD; Mary Ellen Martin, MD; John Meiser, MD; Tuan Nguyen, MD; Vu Nguyen, MD; Neeraja Peterson, MD; Tim Quan, MD; Karen Riedl, MD; Marc Riedl, MD; Kara Rysman, MD; Milene Saavedra, MD; Natalie Sacks, MD; Aaron Shiels, MD; Peggy Shiels, MD; Guido Silversti, MD; Clint Snedegar, MD; Rosa Solorio, MD; Rick Starlin, MD; Steven Stein, MD; Trini Tellez, MD; and Keoki Williams, MD.

RAND Health

Gretchen K. Berland, MD
Leo S. Morales, MD, PhD
Marc N. Elliott, PhD
Jeffrey I. Algazy, MD, MPH
Richard L. Kravitz, MD, MSPH
Michael S. Broder, MD
David E. Kanouse, PhD
Jorge A. Munoz, MS. Phil
Joshua Hauser, MD
Marielena Lara, MD, MPH
Katherine Watkins, MD, MSHS
Hannah Yang
Juan-Antonio Puyol
Linda Escalante
Jennifer Hicks, MPH
Anne Griffin, RN, MPH
Karen Ricci, RN, MPH
Robert H. Brook, MD, ScD
Elizabeth A. McGlynn, PhD

Contents

Executive Summary

Introduction

The Internet is the largest searchable medical library in the world. It is also becoming an increasingly important and influential source of health information for the public. A recent Harris Interactive survey estimated that 97 million adults used the Internet to look for health-related materials. Studies also indicate that consumers are paying attention to what they find; more than 70 percent say online health information has influenced a treatment decision.

Consumers looking for health information on the Internet face a vast array of choices—tens of thousands of health-related Web sites and millions of Web pages of health-related material. How will this wealth of information influence the way that patients interact with physicians and the way that health care delivery is structured?

Optimists predict that ready access to the latest scientific information will allow consumers to participate more actively in their own care, perhaps even eliminating some disparities in access by allowing people to bypass the medical care system. Others worry that the Internet contains incomplete and misleading information, which may directly harm patients by misdirecting a treatment decision. Several studies have documented substantial variability in health-related Web site content for specific diseases. But been no systematic assessments of whether the available Web-based information is sufficiently complete and accurate to support consumer decision making.

Recently several groups have become concerned with ensuring that the health information a consumer finds on the Internet is of high quality. A number of organizations have proposed, published, and voluntarily implemented criteria to guide the evaluation of health-related Web site content. However, these criteria have not been systematically applied to a broad set of Web pages and to multiple medical conditions.

In response to concerns about variation in quality of health information on the Internet, the California HealthCare Foundation asked RAND to design and conduct a large study that described and evaluated English- and Spanish-language health information on the Internet. This study represents a major step forward in rigorous evaluation of Web-based health information.

The study addressed three questions:

- What type of condition-specific information do search engines identify, and how efficient are the engines as tools for locating health information?
- How comprehensive, accurate, and current is the information presented on selected health Web sites?
- What level of literacy is required to understand the information provided by these sites?

The study focused on four medical conditions (breast cancer, childhood asthma, depression, and obesity) in two languages (English and Spanish). These conditions affect diverse populations and are associated with an increased likelihood of early death and disability. Thus they represent the kinds of health care concerns that would motivate consumers to look for information on the Web. A shorter version of this report is also published in the May 23, 2001, issue of the *Journal of the American Medical Association* (JAMA, May 23/30, 2001-Vol.285, No. 20).

Study Design

Evaluating Search Engine Performance. In our assessment of search engine performance, we focused on two basic issues:

1. What are consumers likely to find when they search for specific health topics online? Does the choice of a search engine affect the results of a simple search for health information?
2. How easy is it to find relevant information?

We studied ten English-language and four Spanish-language search engines, selected because of their popularity or their unique method of ranking sites. Using each search engine, we conducted a series of standardized searches using one simple search term for each medical condition. We categorized the results of these searches, including the characteristics of the top Web sites and the type of content found.

Assessing the Quality of Health Information on the Internet. In the second part of the study, we addressed three questions:

1. How comprehensive is the information on selected health Web sites?
2. How accurate is it?
3. How often do Web sites provide documentation that allows one to assess the source or currency of the material?

We selected 18 English-language health Web sites (seven general health and twelve condition-specific) and seven Spanish-language Web sites (three general health and four condition-specific). Six English-language general health Web sites were chosen based on popularity (they were ranked highly in two widely used Internet industry reports by Cyber Dialogue and PC Data Online). Content provided by one of the most popular search engines was also included. We selected

condition-specific English-language Web sites and all Spanish-language Web sites to represent prominent examples of condition-specific sites from commercial, government, and nonprofit educational organizations.

For each of the four medical conditions, we convened panels of three to four nationally recognized clinical experts and representatives from patient advocacy organizations. Each panel developed five to seven essential, "need-to-know" topics and "consumer-oriented" questions, designed to reflect the concerns of patients or family members seeking health information. The questions were very basic, as illustrated by the examples below.

Examples of questions related to treatments for breast cancer:

- "Where can I get information about breast cancer clinical trials?"
- "If I have Stage I or II breast cancer, which is better treatment, mastectomy or lumpectomy plus radiation?"

Examples of questions related to the etiology and symptoms of childhood asthma:

- "What are the common symptoms of asthma in children?"
- "What causes asthma? Is it curable?"

Examples of questions related to the causes and treatments for depression:

- "If my doctor recommends an antidepressant medication for the treatment of my depression, how long should I take it for? What should I expect and when will I feel better?"
- "What causes depression?"

Examples of questions related to treatments for obesity:

- "Should I consider weight-loss drugs, and if so, what prescription drugs are currently available?"
- "Who should consider weight-loss surgery? What are the risks, and how well does it work?"

Based on reviews of the clinical literature, the panels then developed a series of standardized concepts (clinical elements that should be addressed) for the topics and questions. In effect, these concepts represented the kind of clinical information that the panelists thought consumers should find on a site. For example, the clinical elements for depression included the following:

- Antidepressant medications typically begin to work within several weeks. But many patients do not experience substantial benefits for four to eight weeks, and it may take three to four months before patients taking antidepressants feel better.
- Patients with a single episode of acute depression who experience initial improvement should continue to take the medication, usually for six to twelve months after they feel better to keep feeling well.

- With antidepressant medications, many people have some adverse effects early in treatment (in the first four to six weeks); most adverse side effects get better in the first month; for some people, the adverse effects can be bad enough to stop the medicine. Common side effects include anxiety, sexual dysfunction, sleepiness, trouble sleeping, weight gain/loss, restlessness, and nausea.

After the topics and questions were developed, two trained searchers visited each selected Web site looking for information related to the consumer-oriented questions. The results from each search were saved and assembled into separate notebooks. All site identifiers were removed, and the notebooks were sent to physician experts to rate the quality of the material. Four rating forms were developed to help standardize the evaluation.

Readability Assessment of Health Information on the Internet. In the third section of the study, we determined the reading grade level at which information provided by these Web sites is written.

We applied widely accepted readability formulas to randomly selected passages of text from both English- and Spanish-language Web sites. These formulas measure grade levels as a function of the average sentence length and word complexity in text samples.

Key Findings

Search engines are not efficient tools for locating health information on a particular health topic.

- *Few searches lead to relevant health information.* Consumers using English search engines have only a one in five chance of finding information that is relevant to their search. Performance is even poorer for Spanish-language search engines, where consumers have only a one in eight chance of finding relevant information.

- *Search engines take users to different places.* No engine is clearly better than another, but where users start matters. Indeed, different search engines rarely take users to the same site. If the lists of ten Web sites identified by each of two different search engines were compared, only one of the sites would be on both lists.

- *Information on the Internet is commercialized.* A substantial proportion of the information that Internet users are likely to find on Web sites is promotional—i.e., it sells products or services but is not clearly labeled as an advertisement. About half of the information located by English-language search engines is of this nature. For Spanish-language search engines, about one-fifth of the information is promotional.

Consumers often find incomplete answers to important questions, but the information that is provided is generally accurate.

- The average English-language Web site lacked information about one in four of the topics that medical experts and consumer advocates thought were important to consumers. More than minimal coverage was available for only half of the topics.

- Health information on Spanish-language sites was sparse and less consistently accurate. On these sites, half of the four health topics had no coverage at all, and more than minimal coverage was found for only one.

- Nearly two-thirds of the English materials list an author and date. About half of the dated materials were updated within the past year. But only one-sixth of the Spanish materials have dates and authors, and nearly half the materials have neither.

- It is not uncommon for a Web site to contain conflicting information on a clinical topic.

- Coverage varied by topic. For example, breast cancer topics, especially breast cancer screening, were covered significantly more often than all other conditions on English Web sites. In contrast, topic areas related to childhood asthma and obesity were covered significantly less often than the other two conditions on English Web sites. Topics covered least often included symptoms suggestive of poorly controlled asthma and the safety and effectiveness of dietary supplements for treatment of obesity.

- Coverage also varied across Web sites. For example, among the English-language breast-cancer sites, Oncolink provided more than minimal, and completely accurate, coverage significantly more often than the average. Five English sites (oncolink.com, cancernet.nih.gov, cancer.org, onhealth.com, and Webmd.com) provided at least minimal coverage for 70 percent of the breast cancer-related topics for which we searched. Among the depression sites, nimh.gov performed significantly above average, and six English-language sites (nimh.nih.gov, intelihealth.com, Webmd.com, cbshealthwatch.com, onhealth.com, and depression.com) provided at least minimal coverage of 50 percent of the depression-related topics. No site performed statistically better than the average for either childhood asthma or obesity.

Most Web-based health information will be difficult for the average consumer to understand.

According to recent health literacy studies, the majority of many health consumer populations cannot understand material written at the ninth-grade reading level. This means that most health Web sites require reading skills beyond the abilities of many consumers, especially those underserved populations that are most in need of this information.

- Half of the English-language materials are written at the college level, and all were at least the tenth-grade reading level.

- Forty percent of the Spanish-language materials are written at the college level, almost all were written at least the ninth-grade reading level.

Conclusions

Search Engine Performance. More than half of the consumers who use the Internet report that they spend about half an hour looking for health information, so efficiency is an important dimension of search engine performance. Search engines do provide users with an enormous amount of information, but most of it is not relevant to the search. Overall, just one in five links identified by ten English-language search engines and one in eight links from four Spanish-language search engines led to a Web page containing content relevant to the search.

Quality of Health Information on the Internet. Although we retrieved more than 20,000 pages of material related to the "need-to-know" topics and "consumer-oriented" questions, we found substantial gaps

in the availability of key information. For example, less than half of the Spanish-language materials explained that mastectomy and lumpectomy plus radiation are equivalent treatments for early-stage breast cancer. Few sites indicated that a woman with a persistent breast mass and a negative mammogram usually needs further evaluation. If consumers are relying on the Internet to help make important treatment decisions, such deficiencies in information could have serious consequences.

Readability of Health Information on the Internet. Much of the health information available on the Internet is beyond the comprehension of many consumers. All of the English Web site documents assessed had materials that required at least a ninth-grade reading level, and more than half presented material at the college level. Four of seven Spanish-language sites presented materials at the ninth-grade reading level or higher.

Studies of patients in various clinical settings suggest that a ninth-grade reading level is too high for most patients. The reading ability of patients and the readability of health-related information on the Internet must be more clearly matched before the Internet can become an effective medium for patient education or have any potential to reduce disparities in access.

Recommendations

These study findings suggest a variety of steps that consumers, providers, and policymakers could take to help the Internet achieve its promise of making critical health information available to all. Based on our analysis, we make the following preliminary recommendations.

To consumers

1. The Internet health space is unimaginably vast, and navigating through that space is more challenging than you might think. Therefore, it is important to set aside adequate time for a search, and plan to visit several sites. Our searchers found that even finding material for a single consumer question took ten to fifteen minutes per Web site with a high-speed Internet connection.

2. Be aware that sites will not necessarily provide a comprehensive picture of what you need to know about a condition. Web-based information can only supplement consultation with a health care professional. But sites may give you information that helps you ask good questions and understand more clearly what your doctor tells you.

3. Many sites are created and maintained for commercial reasons. This does not mean that the information you find there is inaccurate, but it is wise to consider the possibility that someone is trying to sell you something by providing this information.

4. Do not be surprised if you find information that is conflicting or difficult to understand. Ask a health care provider to help you interpret what you find.

To consumer advocacy groups

1. Advocacy groups could help consumers by "adopting" one or two relevant sites and continually screening their content for coverage and accuracy. One might imagine commercial arrangements between e-health companies (which would produce the

content) and advocacy organizations (which would arrange for impartial review by experts).

2. Advocacy groups could also help consumers by pressing for improvements in site content and presentation to make information more complete, accurate, and accessible, and by referring consumers to good sites.

To health care providers

1. Be aware that the Internet exposes patients to a vast and often confusing array of information, sometimes helping, sometimes hindering, the process of providing good care.

2. Professional societies could play a key role in organizing physicians, pharmacists, or other professionals to provide a more formal interpretive function. In the current health business environment, this may require attention to the mechanisms by which such services could be reimbursed.

3. Specialty societies could provide key clinical content for Web sites. Their involvement could significantly improve coverage, accuracy, and presentation of the material. Specialty societies should work with advocacy organizations to ensure that the material they develop addresses important patient concerns and facilitates patient decision making about when to seek medical care.

To providers of Web-based health information

1. Web site content providers could commission clinical panels of experts to review coverage, accuracy, and factual conflicts before material is put before the public. They could also work with consumer advocacy organizations to make sure that frequently asked consumer questions are addressed and that language used is appropriate for the intended audiences

   Information on consumer-oriented health-related sites should be provided at lower reading grade levels. Currently, the U.S. Department of Health and Human Services recommends that patient education materials not exceed a sixth-grade reading level. A serious effort should be made to reach this standard.

2. Clinical content should be periodically reviewed by experts, and such reviews should be required in all quality assessments of health-related sites. Reviews should be conducted under the auspices of an independent party that is unrelated to the Internet health information provider and does not provide such information itself.

3. Readability standards should also be part of quality assessments of health-related Internet sites. Because readability is an integral part of providing information, explicit standards should be articulated and the methods used for assessing readability should be made public. Readability assessments of materials on Spanish- and English-language Web sites should be made regularly, and results should be widely disseminated in both languages to help consumers choose sites that will best meet their needs.

To policymakers and regulators

1. The best English-language Web sites are far better than the best Spanish-language sites. Major gains in quality of Spanish-language information (and very likely other non-English-language information) can be achieved by translating and culturally adapting what is now available in English.

2. Continue to fund high-quality sites, especially for consumer audiences that are not well served by free market forces alone.

3. Put greater effort into publicizing and increasing access to high-quality government sites.

4. Fund research on how to effectively communicate health information in Web-based format to readers with a wide range of reading levels.

The Internet offers tremendous potential for consumers to easily access important information about their own health problems and those faced by friends and families. This study takes an important first step in conducting rigorous evaluations of Internet medical content. We hope our findings will be used to guide future improvements in the availability and quality of health information on the Internet.

1. Introduction

More than any other medium, the Internet has transformed the way many consumers and health care professionals find health information. Health care is one of the key reasons consumers go online (Cyber Dialogue, 2000). The number of persons who have sought health information online has increased from about 54 million in 1998 to 97 million in 2001, almost a two-fold increase (Taylor, 2001). Studies also indicate that the online population is becoming more representative of the larger U.S. population in terms of race, age, income, and educational attainment. According to several recent surveys, nearly 50% of the Hispanic population reported using the Internet in 2000, and more than half of those reported searching for health information (Taylor, 2001; Fox and Raine, 2000; Fox, Horrigan, Lenhart, et al, 2001).

Consumers looking for health information on the Internet have access to tens of thousands of health-related Web sites and literally millions of Web pages of health-related materials. Because almost anyone can make a Web site available through the Internet, the breadth of information that can be found is vast. On health topics, consumers can readily be linked to scientific information about various health problems and treatments, advertisements for physician and hospital services and pharmaceuticals, advocacy organizations, accounts of personal experiences, and opportunities for people with similar interests to interact online. Studies also indicate that consumers are paying attention to what they find; more than 70% say online health information has influenced a treatment decision (Fox and Raine, 2000).

The rise of consumerism in health care has resulted in increased demand for information. Doctors are interested in patient education because it enhances the ability of patients to participate in their care and may also increase compliance with treatment regimens. But the process of finding high-quality, appropriate health information can be an arduous task. Studies of single medical conditions have suggested deficiencies in the quality of Web-based health information (Sandvik, 1999).

The way that this enormous amount of material will change health care delivery is structured and provided remains to be seen. Optimists predict that ready access to the latest scientific information will allow consumers to make more informed decisions about their health care. Because the information is free, it has been suggested that some disparities in access might be eliminated (United States Bureau of Commerce, 1999). Others worry that the Internet contains incomplete and misleading information, which may directly harm patients by misdirecting a treatment decision

(Ostrom, 1999; Pear, 2000). Another area of recent concern is that of direct access to pharmaceuticals without visiting a physician. There is adequate evidence on the Internet for all of these positions (Armstrong et al, 1999).

Over the past year, policymakers have become more concerned with how to ensure that the information consumers find on the Internet is of high (or at least acceptable) quality (Landro, 2000; Baur and Deering, 2000). Several organizations have proposed and developed guidelines to evaluate health-related Web site content (for example, HON Code, American Medical Association, Internet HealthCare Coalition, Hi-Ethics, MedCertain), but these guidelines have not been systematically applied to a broad set of Web pages and health conditions. Furthermore, little scientific research has been done to empirically assess what a consumer might find while searching on the Internet.

While accessible and high-quality health information on the Internet is important for English speakers, it could be even more useful for Spanish speakers, who often encounter greater barriers to traditional sources of medical care and information. Preliminary data from the 2000 U.S. Census indicate that Hispanics are now the largest minority population in the United States.

In response to these concerns about the variation in quality of health information on the Internet, RAND collaborated with the California HealthCare Foundation to design and implement this project on the evaluation of English and Spanish health information on the Internet. This project represents one of the most comprehensive and systematic efforts ever undertaken in this area. We believe that it is the first project to simultaneously evaluate English-language and Spanish-language search engines and Web sites. The study design addresses three questions:

- What type of condition-specific information do search engines identify, and how efficient are the engines as tools for locating health information?
- How comprehensive, accurate, and current is the information presented on selected health Web sites?
- What level of literacy is required to understand the information provided by these sites?

This study focused on four medical conditions (breast cancer, childhood asthma, depression, and obesity) in two languages (English and Spanish). These are conditions that affect diverse populations and are associated with an increased likelihood of early death and disability. Thus, they represent the kinds of health care concerns that would motivate consumers to look for information on the Web.

Chapters 2 through 4 follow a standard format. We begin by explaining the purpose of this investigation. We then provide a detailed description of the methods employed. In many cases, we had to develop new methods or adapt existing methods to study the Internet. These methods may be useful to others who are interested in evaluating other search engines, other health topics, or Web sites other than those selected for this study. Next we provide the results of our work. We end each chapter with a discussion of the findings.

In Chapter 2 we present our assessment of the performance of 14 frequently visited search engines (ten English-language and four Spanish-language engines). In Chapter 3 we evaluate the comprehensiveness, accuracy, and currency of information presented on general health and condition-specific Web sites. In Chapter 4 we report the reading grade level required to understand information presented on selected Web sites. In Chapter 5 we summarize the results and make recommendations.

2. Study of Search Engine Performance

Purpose

Search engines are tools designed to help people find information on the Internet. More than half of the 100 million individuals who look for health information on the Internet use search engines to do so (Taylor, 2001). Many Internet users search with just a few engines, although more than 3,600 different ones exist (Search Engine Guide, 2001). Most search engines use a combination of methods to identify potentially relevant material: a crawler component (a computer-generated index of Web sites) as well as a directory component (a manually edited index of Web sites) (Search Engine Watch, 2001). After identifying material, search engines employ different methods to rank the sites they have found. Examples of ranking methods include:

- Ranking by the location and frequency of key words within a Web site
- Ranking by the number of times a site is linked to by other Web sites
- Ranking by payment from Web sites
- Ranking by human editing

In our assessment of search engine performance, we focused on two basic issues. Specifically, we wanted to know:

- What are Internet users likely to find when they search for specific health topics online? Does the choice of a search engine affect the results of a simple search for health information?
- How easy is it to find relevant information? Does this vary by health condition?

Methods

Because this study is one of the first to attempt to characterize the way in which search engines operate to help consumers find online health information, we designed a structured evaluation of each medical condition rather than observing the experiences of typical users. This section describes

how we selected the 14 search engines included in the study, the methods that were used to evaluate the performance of those search engines, and the analytic techniques used to assess the results. Figure 2.1 illustrates the flow of the search engine study.

Selecting Search Engines. Working with three search engine experts,[1] ten English-language and four Spanish-language search engines were selected for this part of the study. We chose search engines based on either their popularity (number of unique visitors per month as reported by Media Metrix, Inc. in June 2000) or based on the method by which the search engine ranked Web sites. Three of the English-language and two of the Spanish-language search engines were chosen based on popularity. The remaining nine search engines were selected because they featured unique ranking methods of ranking Web sites.

The ten English-language search engines selected were AltaVista, Ask Jeeves, Direct Hit, Excite, Google, Go To, Lycos, Metacrawler, Northern Light, and Yahoo! The four Spanish-language search engines selected were: Quepasa, TeRespondo, Yahoo! Espanol, and Yupi. Tables 2.1 and 2.2 describe some of the characteristics of these search engines.

Training Searchers and Coders

The first phase of our study involved searching the selected search engines for links addressing the selected medical conditions and then categorizing those pages found by the relevance of their content. To do so, five experienced computer users with college degrees but no medical training served as searchers for the English-language search engines. All had some previous experience using computers and the Internet. Two fluent Spanish-speakers from UCLA and RAND served as the searchers for the Spanish-language search engines.

During a full-day training session using CatchtheWeb® Software (described in detail below), searchers learned how to define and identify relevant links and save this information into a database. An inter-rater reliability test was performed for each independent task that the searchers were required to perform. The study did not begin until the inter-rater reliability of the group was acceptable (defined as a kappa statistic of greater than 0.8 for each task).

Four coders were then trained to systematically categorize the relevance of content on the Web pages that had been saved by the searchers. The coders for the English materials were health-services researchers with graduate training in public health. The coders for the Spanish materials were fluent Spanish-speakers with graduate degrees. A full-day training session was conducted, including multiple standardized exercises designed to train coders to recognize content related to the four conditions of interest. An inter-rater reliability test was performed for each independent task that the coders were required to perform. Classification of the content pages did not begin until the inter-rater reliability of the group was adequate (defined as a kappa statistic of greater than 0.8 for each task).

[1] Danny Sullivan (editor of Search Engine Watch.com), Tamas Dosckocs (National Library of Medicine), and Dan Durazo (Durazo Communications)

Figure 2.1: Flow of Search Engine Study

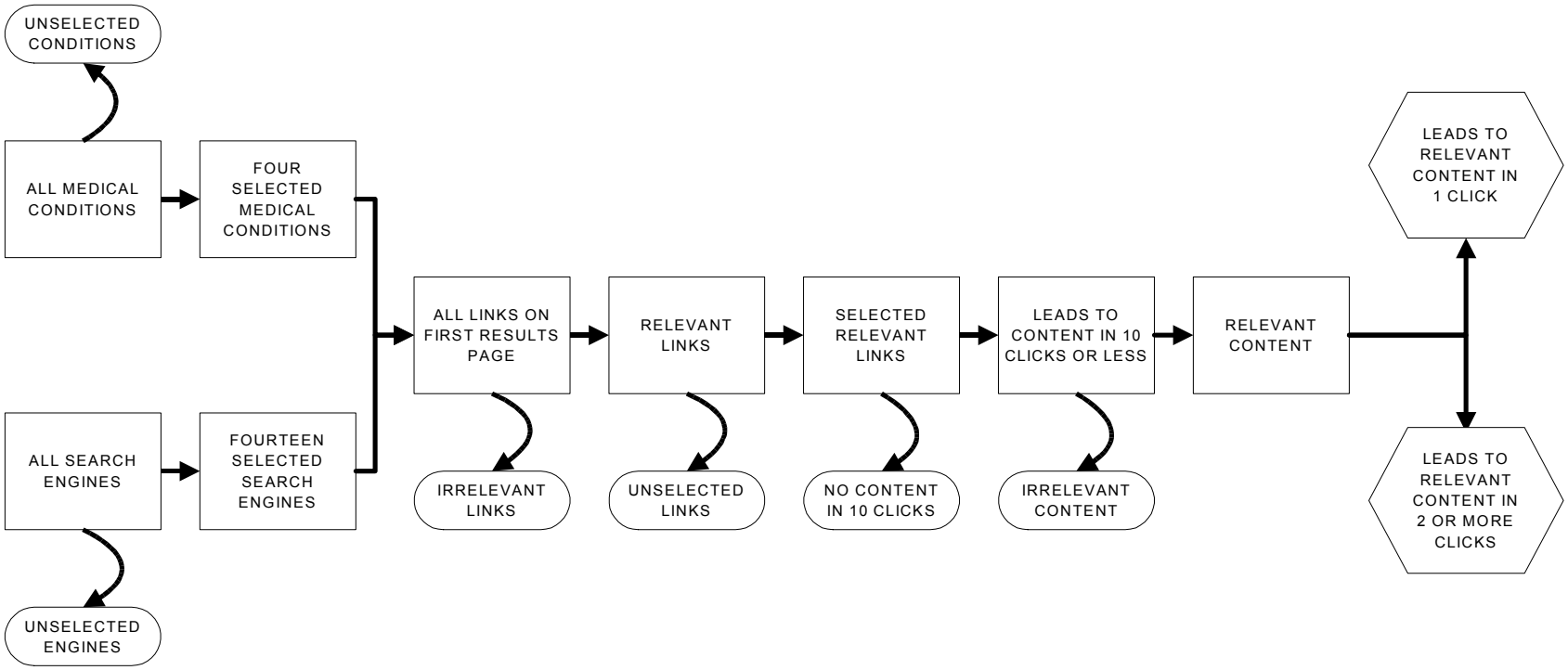


Table 2.1: Search Engines – Types and Characteristics

SEARCH ENGINE	SEARCH ENGINE TYPE	REASON SELECTED	CRAWLER CHARACTERISTICS	DIRECTORY CHARACTERISTICS
AltaVista	Crawler; directory component	Crawler results ranked based on how many times search words appear in the text.	Full text search HTML title First lines of text Order of appearance "Sponsored listings" section (Go To) No submission fee	LookSmart directory Open directory Submission fees: Express: $199 Basic: $99 Not-for-profit: free
Ask Jeeves	Directory; metacrawler component	Human-edited directory results presented in the form of a question.	Metacrawler: AltaVista, Excite, WebCrawler, Go To	Editors and popularity technology Editors identify Web sites Pick "best in class" content No submission fee
Direct Hit	Directory; crawler component	Crawler results ranked based on the number of times users click on the site link at partner Web sites.	Previous site activity determines rank Human input No submission fee	Popularity engine Open directory No submission fee
Excite	Crawler; directory component	Ranked #3 by Media Metrix, June 2000.	Page content Meta-tags Referring anchor text Link popularity No submission fee	LookSmart directory Submission fees: Express: $199 Basic: $99 Not-for-profit: free

Table 2.1: Search Engines – Types and Characteristics

SEARCH ENGINE	SEARCH ENGINE TYPE	REASON SELECTED	CRAWLER CHARACTERISTICS	DIRECTORY CHARACTERISTICS
Google	Crawler; directory component	Crawler results ranked based on the number of times other sites are linked to the ranked site.	Page relevance Proximity of search terms on page Popularity Text matching "PageRank" No submission fee	Open directory No submission fee
Go To	Directory; crawler component	Directory results ranked based on how much the Web site author is willing to pay.	Inktomi crawler	Editor determined relevance Paid placement Advertisers bid in ongoing auction Search results ranked in order of price
Lycos	Directory; crawler component	Ranked #2 by Media Metrix, June 2000.	FAST crawler Inktomi crawler No submission fee	Open directory No submission fee
Metacrawler	Metasearch engine crawler; directory component	Presents the results of several search engines ranked by how many search engines list the Web site.	No local database Queries other search engines Presents "voted" order	LookSmart directory Open directory

Table 2.1: Search Engines – Types and Characteristics

SEARCH ENGINE	SEARCH ENGINE TYPE	REASON SELECTED	CRAWLER CHARACTERISTICS	DIRECTORY CHARACTERISTICS
Northern Light	Crawler	Crawler results ranked based on patented classification system and relevance rating.	Natural language search engine Frequency of search terms in text Context of search terms in page Link popularity No submission fee	N/A
Yahoo!	Directory; crawler component	Ranked #1 by Media Metrix, June 2000.	Google crawler	Yahoo! editors "suggest your site" Relevance Submission fee: -Free for non-commercial -$199 for commercial -$600 for adult content
Yupi	Directory; crawler component	Identified as Popular Spanish Search Engine by search engine experts, and unique directory.	AltaVista Crawler	Manual selection Manual categorization Human editors
QuePasa	Crawler	Identified as Popular Spanish Search Engine by search engine experts and only Spanish Inkomi-powered search engine.	Inktomi-powered Link and caching data to determine popularity Link analysis Paid inclusion	N/A

Table 2.1: Search Engines – Types and Characteristics

SEARCH ENGINE	SEARCH ENGINE TYPE	REASON SELECTED	CRAWLER CHARACTERISTICS	DIRECTORY CHARACTERISTICS
TeRespondo	Directory; metacrawler component	Identified as Popular Spanish Search Engine by search engine experts and only Metacrawler.	Queries other search engines Yahoo! en Espanol, AltaVista, HotBot, Excite, Fast	Editor-controlled Editors generate a list of questions leading to advertisers' sites
Yahoo! en Español	Directory; Crawler component	Identified as Popular Spanish Search Engine by search engine experts and only Google-powered search engine.	Google crawler	Yahoo! editors "suggest your site" Relevance Submission fee: Free for non-commercial $199 for commercial ($600 for adult content or services)

Crawler = computer-generated index of Web pages
Directory = human-generated index of Web pages
Open Directory = comprehensive human-generated index of Web pages run by volunteer editors
Inktomi = proprietary search technology for search engines by Inktomi, Inc.

Table 2.2: Search Engines – Revenues and Advertising

SEARCH ENGINE	COMPANY INFORMATION	INFO ON SITE	ADVERTISING DETAILS
AltaVista	Majority-owned company of CMGI, Inc. Publicly held NYSE: CMGI Revenues quarter ending 10/31/00: $366.1M Search and Portals business unit (AltaVista, iCAST, MyWay.com) Revenues quarter ending 10/31/00: $98M No specific information on generation of revenues or advertising revenues	Contact for info	**Per "DoubleClick" Advertising Site:** **Keyword Targeting**: allows advertisers to tie their banners to particular search terms (keywords) and therefore target their message to the appropriate audience **Categories**: targets users who search by subject, with banners and sponsor tiles tailored to meet advertisers' objectives **Sponsorships**: delivers high-profile positioning on homepages and results pages, for enhanced visibility and brand recognition (available for AltaVista, AltaVista/Subject Search categories as well as AltaVista Channels) **General Rotation**: untargeted banners that appear on AltaVista's results pages **Home Page Badge**: (230x33) can be targeted by geography, browser and time of day **Run of Site**: untargeted banners delivered throughout AltaVista **AltaVista Channels**: provide dynamic links to related content based on user's keywords and Navigation Bar tabs. Banners may be targeted to users of the following content: Live, Money, News, Sports, Travel, Career, Health, Entertainment, World, Technology, Women, and Real Estate.
Ask Jeeves	Publicly held NASDAQNM: ASKJ Revenues quarter ending 9/30/00: $29M Web Properties Group (Ask.com, DirectHit.com, AJKids.com) Revenues quarter ending 9/30/00: $15.8M Revenues generated from advertising, e-commerce, licensing to corporations Advertising revenues quarter ending 9/30/00: $8M	Some pricing info present	**Text Sponsorship**: bid for search topics; pay only when link is shown **Answer Link**: editors assist in crafting questions; "answer" is given at advertisers' site by direct link; Ask Jeeves leads user to further "explore" site **Content Modules**: links (headline, text description, graphics) along right of page within topic area **Advertising Services**: banners/buttons triggered by keywords or categories **Dynamic Response**: similar to answer link, but advertiser messaging varies with user query **Shopping Guide**: users interested in specific products are sent directly to a point of purchase on advertiser's site
Direct Hit	Subsidiary of Ask Jeeves	Some pricing info present	**Text Sponsorship**: bid for search topics on Ask Jeeves; pay only when link is shown **Advertising Services**: banners/buttons triggered by keywords or categories
Excite	Publicly held At Home Corporation	Present	**General Rotation**

Table 2.2: Search Engines – Revenues and Advertising

SEARCH ENGINE	COMPANY INFORMATION	INFO ON SITE	ADVERTISING DETAILS
	NASDAQ: ATHM Revenues quarter ending 9/30/00: $160.5M Revenues generated from media and advertising, subscriber network and services, international Advertising revenues quarter ending 9/30/00: $77.7M		**Channel Rotations** **Individual Keywords** **Keyword Phrases**
Google	Privately held Venture Capital Firms ($25M – 6/99): Sequoia Capital, Kleiner, Perkins, Caulfield & Byers Other inv: Stanford Univ., Andy Bechtolsheim, Ram Shriram	Present	**Premium sponsorship** (top of page): Keyword and keyword phrasing : text-based ad appears at the top of a results page whenever the keyword purchased is included in user's search. Categories: advertisers can select the categories most appropriate to their business and Google will match the most relevant category ads to each user's search **AdWords Program –** for site owners (right side of page): Keywords: search must contain entire word Negative keywords: allows for elimination of unrelated keywords Phrase Matches: queries must match the words in the phrase exactly
Go To	Publicly held NASDAQ: GOTO Revenues quarter ending 9/30/00: $25.1M Revenues generated from search listing advertisements (90%) and banner ads (10%) Advertising revenues quarter ending 9/30/00: $25.1M	Present	**Keyword:** bid for relevant keywords No information on banner advertising
Lycos	Publicly Held NASDAQ: LCOS Revenues quarter ending 7/31/00: $88M Terra Lycos (TRLY) as of 10/31/00 Revenues generated from advertising, e-commerce and other Advertising revenues quarter ending 7/31/00: $56.9M	Present	**Banners** **Keyword Advertising** **Key Phrase Advertising** **Shopping Network** **Travel Network** **Lycos Community Offerings** Run of service advertising for Chat, Email, Personal Guide. **Targeted Navigational Services:** Road maps, people finder, etc.

Table 2.2: Search Engines – Revenues and Advertising

SEARCH ENGINE	COMPANY INFORMATION	INFO ON SITE	ADVERTISING DETAILS
Metacrawler	Infospace Publicly held NASDAQ: INSP Revenues quarter ending 9/30/00: $57.7M Consumer services revenues generated from advertising, subscriber fees, and transaction fees No specific information on advertising revenues	Present	**Banners, Buttons** **MetaCrawler Results Sidebar** **MetaCrawler Suggests** **Keyword Banner** **Shopping Partners** **Keyword Textlink** **Jukebox** **Audio Sidebar Button** **Keyword Targeted Banners** **Favorite Searches Sponsorship**
Northern Light	Privately held Investors: Reuters Group PLC, Hewlett-Packard, Times Mirror Co.	Present	**Run of Site:** ad banners at the top of the pages are displayed in rotation across all advertising pages **Site Specific:** ad is displayed on specific, highly trafficked sections of the site **Category** **Keywords and phrases** **Keyword Text Links:** each link can be keyword coded, allowing textual information to pop up alongside the search results for that word **Skyscraper Ads:** larger ad space to convey a complex message, technical topic, etc. **E-mail News Alerts:** daily e-mail on news stories in customer's interest area **Special Edition / Micro Site:** "micro-site" offering unbiased third-party view of information resources and live updates of particular interest to a target audience **Sponsorships:** customized marketing solutions
Yahoo!	Publicly held NASDAQ:YHOO Revenues quarter ending 9/30/00: $295.5M Revenues generated from banner and sponsorship advertisements, business services, e-commerce, barter transactions Advertising revenues quarter ending 9/30/00: $265.9M	None, e-mail sales	Advertising information from Yahoo Fusion Marketing site: **Banners** **Buttons/Text Links** **Event Marketing (Online and Offline)** **Sponsorships** **Promotions (Front Page, Custom, Multi-Sponsor, etc.)** **Streaming Ads (Audio and Video)** **Enhanced Ad Units**

Table 2.2: Search Engines – Revenues and Advertising

SEARCH ENGINE	COMPANY INFORMATION	INFO ON SITE	ADVERTISING DETAILS
Yupi	Privately held No specific information on generation of revenues or advertising revenues	Present	Preferential listing in Yupi categories Individual keywords Focused impressions
QuePasa	Publicly held NASDAQ: PASA Revenues quarter ending 9/31/00: $966,581 Major Investors: Telemundo Revenues generated from banner and sponsorship advertisements, e-commerce, strategic agreements with other e-services companies	Contact for info	**Banners** **Sponsorships**
TeRespondo	Privately held No specific information on generation of revenues or advertising revenues	Present, contact for pricing	**Answer Link**: editors assist in crafting questions; "answer" is given at advertisers' site by direct link; TeRespondo leads user to further "explore" site
Yahoo! en Español	Publicly held NASDAQ:YHOO Revenues quarter ending 9/30/00: $295.5M Advertising revenues for quarter ending 9/30/00: $265.9M Revenues generated from banner and sponsorship advertisements, business services, ecommerce, barter transactions	None, e-mail sales	Advertising information from Yahoo Fusion Marketing site: **Banners** **Buttons/Text Links** **Event Marketing (Online and Offline)** **Sponsorships** **Promotions (Front Page, Custom, Multi-Sponsor, etc.)** **Streaming Ads (Audio And Video)** **Enhanced Ad Units**

CPM = cost per thousand impressions
CPC = cost per click

Evaluating Characteristics of Top Web Sites Listed by Search Engines

To answer our first research question about search engines, we compared the extent to which search engines identify the same or different Web sites (measured as degree of overlap), when searchers entered four simple search terms into the query boxes of each of the 14 search engines studied. The search terms were:

- breast cancer
- childhood asthma
- depression
- obesity

Because prior research suggests that individuals searching online tend to look at the Web sites listed first following a search, we compared only the degree of overlap among the first ten Web sites identified by each of the search engines (Cyber Dialogue, 2001).

Comparing the Efficiency of Search Engines

To answer our second research question about search engines, we examined the efficiency of search engines in two ways. First we compared the proportion of relevant links on the first page of results on each search engine. Second, we compared the proportion of relevant content pages obtained by following ten relevant links on the first page of results on each search engine. These comparisons were based on the search engine results obtained by entering the same search terms described in the preceding paragraph.

We defined a link as relevant if (1) the search term of interest (i.e., breast cancer, childhood asthma, depression, or obesity) was present in the link itself or in the surrounding text or (2) one of 30-40 related key terms (i.e., tamoxifen, inhaler, gastric bypass surgery, St. John's wort) was present in the link itself or in surrounding text. The list of related terms created for the English-language search engines was modified for cultural appropriateness and translated for use on the Spanish-language search engines. The lists are provided in Appendix A.

After all relevant links on the first page of each search engine's list were identified, searchers then followed a sample of ten relevant links to determine whether they led to relevant content. The ten relevant links were chosen as follows:

- the first five consecutive relevant links on the search results page
- and a stratified random sample of the remaining links. (All remaining links were enumerated and assigned to one of five strata of equal size according to order. Then one link from each strata was randomly selected).

Searchers clicked on the selected relevant links until they reached a Web page with content (defined as when 50% of the space occupied contained text that was not primarily an index of the Web site). If the first relevant link led to a content page, they saved the page for further analysis. If the first relevant link led to another list of links, the searcher identified the first 15 relevant links on that page and then selected one of those links to follow. From that list of 15 relevant links, a random one of

those was selected for the searcher to follow. This process was repeated up to ten times. If after ten of these cycles, searchers had not reached a content page, the last page was saved for analysis.

Characterizing the Content on Web Pages. Trained coders used a rating form to characterize the content on the Web pages identified by the search study described in the preceding paragraph. The content on each Web page was classified according to: (1) relevance of the identified content; (2) type of medical content (e.g, alternative, allopathic); and (3) number of advertisements. Web pages were considered to be relevant if they had any material related to the search terms (breast cancer, childhood asthma, depression, and obesity). Sponsorship was classified by type of site providing the relevant content (for example, advocacy, medical organization, e-health). Web pages that sold information, services, or goods were classified as having promotional material (defined as material designed to encourage site visitors to purchase products or services or participate in research programs sponsored by the site). Advertisements were classified separately from promotional material and defined as advertising material only if physically located in one of two specific positions (banner or sidebar) on the Web page.

Data Collection and Management

The study was conducted between September 25 and September 28, 2000. Working in a RAND computer laboratory, searchers used Dell Pentium II Processor Computers equipped with Sony Multiscan 100sf monitors and S3 ViRGE-DX/GX Video Cards. Each system was configured with 64 megabytes of RAM and connected to the RAND Internet by an Ethernet Connection. All computers were operated using Windows 98 and Internet Explorer 5.0 with cookie files (used by Web sites to identify when Internet searchers visit their sites) disabled.

One of the methodological challenges in studying health information on the Internet is the ability to study the Web pages as they appear on the user's screen. For this study, we used an Internet software application called Catch theWeb® produced by Math Strategies in Greensboro, North Carolina. This software enabled the project researchers to accurately save the Web pages for use at a later date as they appeared on the screens. This ensured that coders classified the pages as they would have appeared to a user on the screen.

Analytic Methods

All analyses were conducted separately for English and Spanish sites and search engines. All statistical tests were two-sided, and were assessed for significance at the 0.05 level.

The unit of analysis was the link (specific URL [uniform resource locator]). For analyses of Web site overlap and the proportion of links leading to relevant content, the universe was all selected links. For analyses of the proportion of links that were relevant, the universe was all links on the first results page.

For each relevant link, the number of engines listing that link in their top ten relevant links was noted. The uniqueness of a search engine's links was measured by computing the average number of other engines finding each of that engine's top relevant links, then dividing by the number of other engines. This yielded the probability that a single randomly selected search engine would list a randomly selected link from the first engine's top ten links in its own top ten list of links.

In evaluating the proportion of links leading to a given outcome, an omnibus 10 x 2 (9 degree of freedom) chi-squared test of homogeneity was used to assess whether proportions differed among

the search engines. If the omnibus test was significant, a series of ten 2 x 2 (1 degree of freedom) chi-squared tests contrasting the proportion in each search engine to the mean proportion of the remaining search engines was conducted. Search engines yielding statistical significance in this series of tests were considered significantly different from the overall mean. A similar two-stage procedure was used to assess the existence of differences in proportions by condition.

Results

Degree of Overlap. Figure 2.2 shows the degree of overlap among English-language search engines in the top ten Web sites identified in response to a structured search. On average, 11% of Web sites found by one English search engine appeared on the top ten list of another search engine (range 1-24%). For Spanish-language search engines, the average was 25% (range 11-33%) (Figure 2.3).

Efficiency of Searches and Relevance of Content by Search Engine. The first page of search results from all ten English-language search engines listed 3,735 links, 1,265 (34%) of which were relevant. A typical search produced a list of 93 links of the first page of search results, about one-third of which were relevant to the search (had the search term or a related word in the title). Among the English-language search engines studied, AltaVista, Direct Hit, and Metacrawler produced higher-than-average proportions of relevant links, while Excite and Northern Light produced lower proportions. On average, three in five of these relevant links reached information related to the search. Relevant links found using Northern Light and Google were significantly more likely to reach information related to the search; relevant links using Direct Hit, Go To, and AltaVista were significantly less likely to do so. Relevant links led to relevant content within ten clicks 59% of the time (Table 2.3, column 3). Overall, consumers using English search engines have a one in five chance of finding information that is relevant to their search.

To address the concern that relevant links sampled from the lower portions of the first results page might have been less likely to lead to relevant content, we compared the first five relevant links on the first results pages to other sampled relevant links. Sixty-one percent of the former links and 58% of the latter links led to relevant content, a difference that was not statistically significant. Overall, relevant content was one click away from a relevant link about one-third of the time with considerable variation among search engines (range, 10-52%) (Table 2.3, column 5).

The first page of search results from all Spanish-language search engines listed 1,685 links, 296 (18%) of which were relevant. A typical search produced a list of 105 links, fewer than one-fifth of which were relevant. Yahoo! Espanol produced higher-than-average proportions of relevant links, while Yupi produced lower proportions. On average, about three in five of these relevant links reached information related to the search. Relevant links found using TeRespondo were significantly more likely to reach information related to the search; relevant links using Quepasa were significantly less likely to do so. When links led to information related to the search, it required more than one click to find this information 62% of the time.

Like the English-language search engines, the first five relevant links on the first results page were no more likely to lead to relevant content than other sampled relevant links (58% and 69%, respectively, reached relevant content, a difference that was not statistically significant).

Efficiency of Search Engines by Condition. The proportion of links identified as relevant on English search engines varied by condition with obesity having the lowest performance (23% of links were relevant) and breast cancer the highest (46% of the links were relevant), as shown in Table 2.4.

Figure 2.2: Amount of Overlap Among Search Results*
English Search Engines

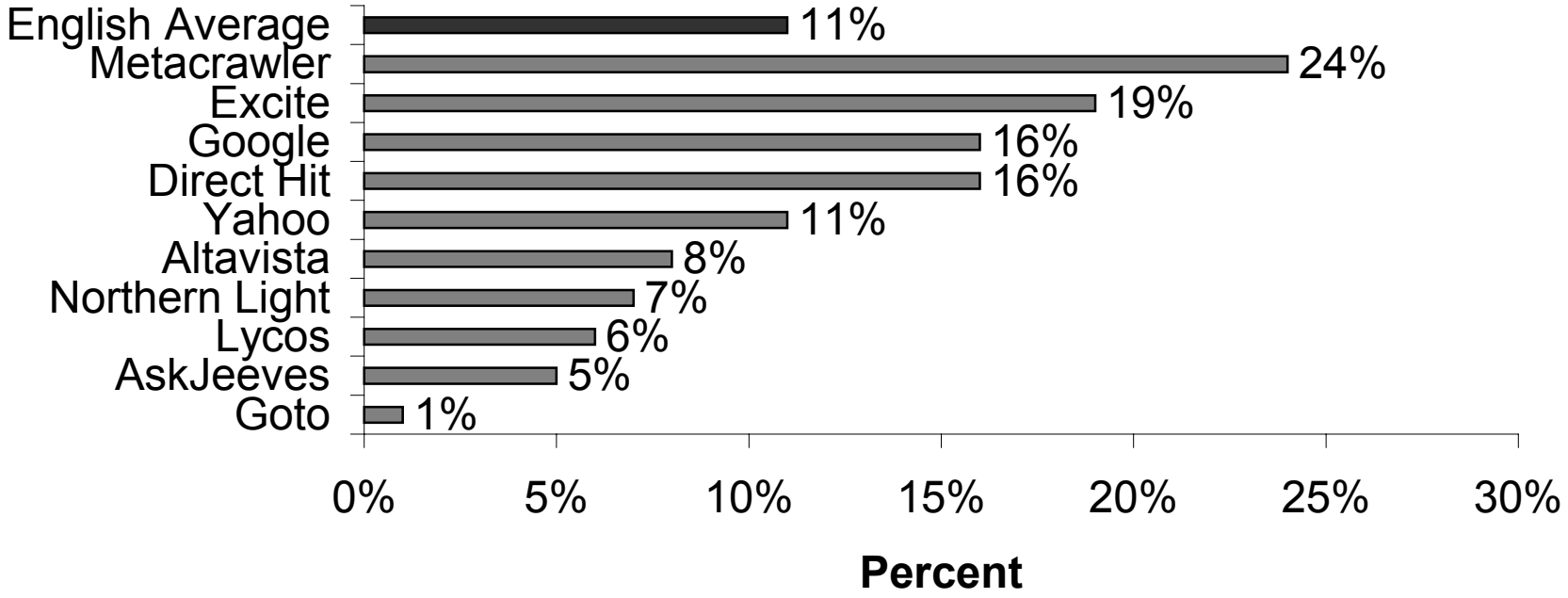
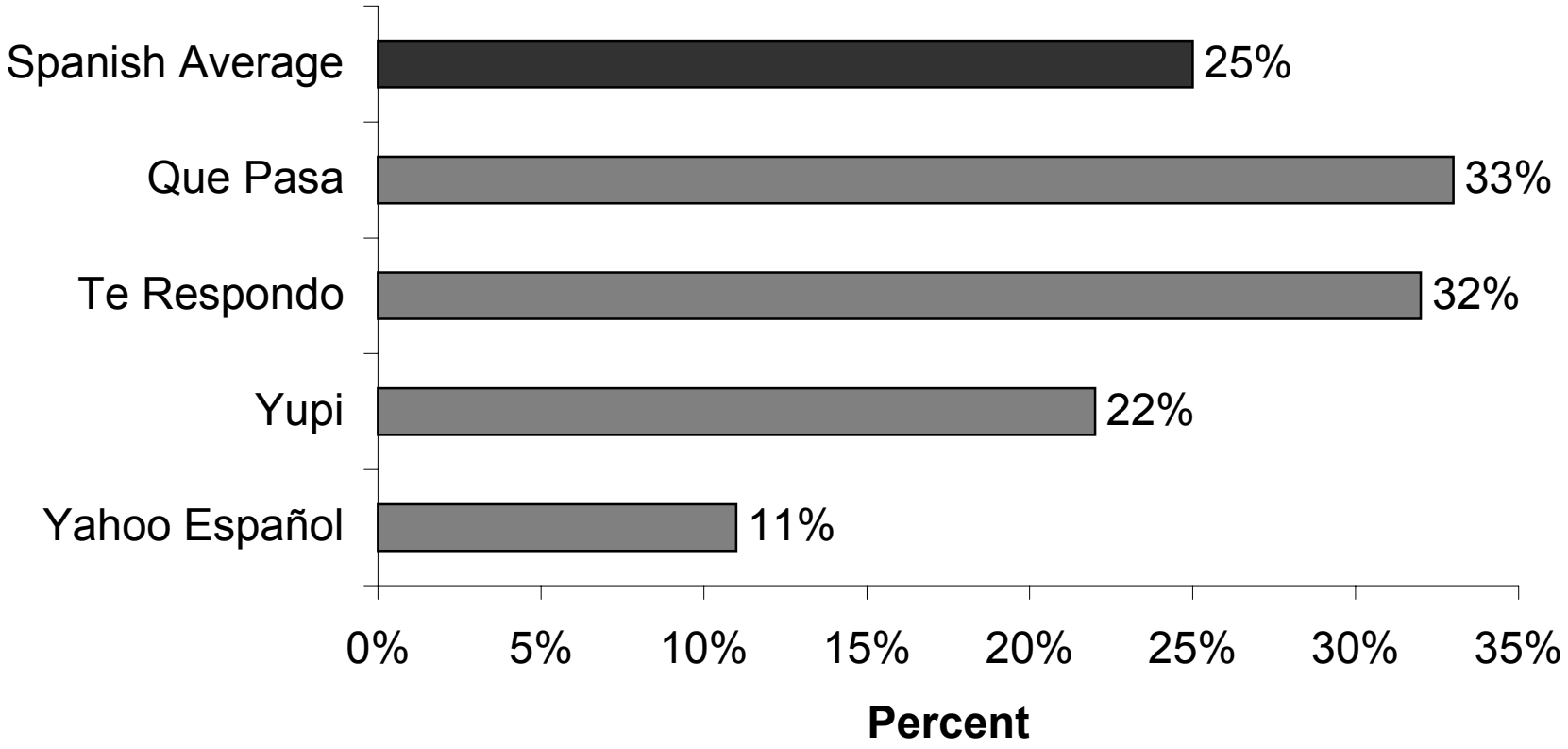


Figure 2.3: Amount of Overlap Among Search Results*
Spanish Search Engines



*Proportion of overlap of top ten Web site lists for two randomly selected English-language search engines.

Table 2.3: Efficiency of Reaching Relevant Content, by Search Engine

Search Engine	Total Number of Links on First Results Page (Four Medical Conditions)	Number of Relevant Links from First Results Page (%)	Percentage of Pages with Relevant Content Reached from Selected Relevant Links**	Percentage of Pages with Relevant Content Reached in One Click from Selected Relevant Links**
English Engines	**3735**	**1265 (34%)**	**59**	**31**
AltaVista	282	142 (50%)	35--	22
Ask Jeeves	1198	299 (25%) --	68	48++
Direct Hit	302	129(43%) ++	38--	10--
Excite	232	61 (26%) --	60	47++
Google	163	51 (31%)	74++	23
Go To	344	155 (45%) ++	43--	25
Lycos	353	132 (37%)	50	18
Metacrawler	384	153 (40%) ++	70	22
Northern Light	241	55 (23%) --	88++	43
Yahoo!	236	88 (37%)	67	52++
Spanish Engines	**1685**	**296 (18%)**	**63**	**38**
QuePasa	283	52 (18%)	59	18--
Te Respondo	853	105 (12%) --	78++	60++
Yahoo! en Espanol	181	63 (35%) ++	66	45
Yupi	368	76 (21%)	49--	28

++Statistically better performance than the mean of other search engines within language ($p<=0.05$)
--Statistically worse performance than the mean of other search engines within language ($p<=0.05$)
**Assuming randomly selected relevant links are representative of all relevant links

Table 2.4: Efficiency of Reaching Relevant Content, by Condition

Condition	Total Number of Links on First Results Page (4 Medical Conditions)	Number of Relevant Links from First Results Page (%)	Percentage of Pages with Relevant Content Reached from Selected Relevant Links**	Percentage of Pages with Relevant Content Reached in One Click from Selected Relevant Links**
English Engines	**3735**	**1265 (34%)**	**59**	**39**
Breast Cancer	802	366 (46%) ++	53	24
Childhood Asthma	738	263 (36%)	69	48 ++
Depression	882	328 (37%) ++	57	30
Obesity	1313	308 (23%) --	59	22 --
Spanish Engines	**1685**	**296 (18%)**	**63**	**38**
Breast Cancer	419	76 (18%)	58	45
Childhood Asthma	402	66 (16%)	79	41
Depression	429	62 (14%)	56	31
Obesity	435	92 (21%)	60	35

++Statistically better performance than the mean of other conditions within language ($p<=0.05$)
--Statistically worse performance than the mean of other conditions within language ($p<=0.05$)

Childhood asthma had the highest proportion of relevant content pages reached from relevant links in one click (48%). Obesity had the lowest proportion of relevant content pages reached from relevant links in one click (22%).

On the Spanish-language search engines, the proportion of relevant links was low, ranging between 14% and 21% for the four medical conditions. The proportion of pages with relevant content reached from relevant links ranged from 79% for childhood asthma to 58% for both breast cancer and depression. The proportion of relevant content pages reached in one click from a relevant link was also highest for childhood asthma (62%) and lowest for obesity (38%).

Type of Information Found. Allopathic medical content (for example, information on surgical options for breast cancer) was found on two-thirds of English-language relevant content pages. Alternative medical content (for example, information on herbal therapies for breast cancer) was found on 11% of English-language relevant content pages. Results in Spanish were similar to English.

Commercialization. Both explicit advertisements (banner and sidebar ads) and promotional material (products or information promoted outside of the banner or sidebar position) were common on relevant content pages. More than half of relevant content pages contained explicit advertisements and 44% contained promotional material that was not presented as an explicit advertisement. Pages without alternative medical content were less likely to contain promotional material than pages with such content (32% vs. 64%, $p<0.05$). The presence of advertisements and promotional materials on relevant Spanish-language content pages was 36% and 21%, respectively.

Discussion

The tremendous amount of information available to consumers is clearly one of the major attractions of the Internet as a means for obtaining health information, but consumers must sift through a large amount of materials during their searches. This study found that search engines are only moderately efficient in locating information on a particular health topic and the efficiency with which relevant information can be found varies significantly across search engines and conditions. Overall, just one in five links identified by ten English-language search engines and one in eight links from four Spanish-language search engines led to a Web page with content relevant to the search. More than half of consumers who use the Internet report that they use search engines to find health information, and they spend about a half-hour on such searches (Carolyn Gratzer, Cyber Dialogue, Oral Communication, October 13, 2000), so efficiency and the relevance of information retrieved are important aspects of performance.

Although advertising and other non-explicit promotional material were common, it was beyond the scope of this study to evaluate whether or not consumers have difficulty recognizing commercial or promotional health information.

In addition, when we reviewed the top ten Web sites listed on each of the search engines, the degree of overlap was quite small (only 11% overall). This level of variability across search engines and conditions suggests that the likelihood of finding the information one needs varies considerably depending on which search engine is used. No search engine is clearly better than another, but where users start matters.

This section of the study has some important limitations worth noting. First, the Internet changes constantly, and we were able to study it at only one point in time. But without concerted attention to

the issue, it seems unlikely that the variability in performance is likely to change dramatically. Second, we looked at a small sample of search engines and conditions, and hence we cannot draw more general conclusions about the performance of all search engines and information on all conditions. But because we included the most popular search engines, the results should reflect what most people experience. Third, we studied the performance of search engines using very simple search terms describing the medical condition. Our findings regarding the efficiency of search engines in yielding relevant content might have been quite different if more sophisticated search strategies were employed. Fourth, the research conducted here was not a naturalistic experiment (for example, using actual consumers to search for information and testing their knowledge after such a search) so we cannot draw conclusions about what consumers actually encounter when they search for information, or about how well they are able to judge the quality of the information they find. But the systematic nature of our methods provides a backdrop for future studies of actual consumer behavior—we can compare what consumers are able to find with what is actually out there to find.

3. Quality of Health Information on the Internet

Purpose

Today, the millions of individuals who go online to find health information can choose from tens of thousands of health-related Web sites, each offering pages and pages of materials on health. Unlike some of the traditional approaches to obtaining health information, they are accessible to anyone with a computer at any hour, and most are free. When individuals go online, surveys indicate many are looking for information on specific medical conditions (Cyber Dialogue 2000). Obesity and cancer were among the most heavily researched medical conditions (Cyber Dialogue 2000). Studies also show that Internet users pay attention to what they find online. More than 70% say online health information has influenced a decision about their treatment (Fox, 2001).

To date, little research has been conducted to try to describe the types and quality of information an individual might find on a variety of English- and Spanish-language health-related Web sites for multiple health conditions. In this chapter, we explore three questions about health information on the Internet:

- Do Web sites provide information related to key clinical questions for each of the four medical conditions studied (breast cancer, childhood asthma, depression, and obesity)?
- If information is provided on these topics, is it accurate?
- How often do Web sites voluntarily provide information related to rating quality criteria such as authorship and currency?

Methods

To answer the questions posed in this chapter we selected four target medical conditions, chose 26 English- and Spanish-language Web sites (both general health and condition-specific), developed specific questions on which Web sites were evaluated for coverage and accuracy, selected material relevant to those questions on each Web site, conducted a systematic review of the material, and analyzed the results. Each of these steps is described below in greater detail.

Selecting Medical Conditions

We selected four target medical conditions affecting diverse populations: breast cancer, childhood asthma, depression, and obesity. These conditions were chosen because they are prevalent, affect diverse populations, and represent conditions for which many consumers might seek online information (Allison et al, 1999; Mokdad et al, 1999; Mannino et al, 1998; Kessler, 1994). The conditions cause significant morbidity and mortality, so health information may have the potential to greatly improve patient education and participation in the management of their health problems. While the findings for these four conditions cannot be generalized to all health information on Web sites, they provide a broad overview of what consumers are likely to find.

Selection of Web Sites

Eighteen unique English-language and 7 unique Spanish-language health Web sites were selected for this study; each provided information on some or all of the four medical conditions. Web sites were chosen based on either popularity or because they were devoted to one of the conditions we were studying. We selected six English-language general health Web sites that were ranked highly in two widely used Internet industry reports, Cyber Dialogue and PC Data Online for September 2000. Content provided by one of the most popular search engines was also included. Sites providing information on more than one condition are referred to as general consumer health Web sites; sites providing information on only one condition are referred to as condition-specific Web sites. Condition-specific English-language Web sites and all Spanish-language Web sites were selected by project staff to represent prominent examples of condition-specific Web sites from commercial, government, and nonprofit educational organizations. The Spanish-language Web sites included four condition-specific sites and three general consumer sites. The English- and Spanish-language Web sites we studied and their specific URL addresses are listed in Table 3.1.

Background Information on English- and Spanish-Language Web Sites

There has been considerable attention in the press about the financial performance of a variety of companies on the Internet. In Tables 3.2 and 3.3 we provide some descriptive information about the English- and Spanish-language Web sites we studied, including revenue sources, nature of the advisory board (if present), and who is responsible for writing the content offered on the Web site.

English-language Web sites. Of the English-language Web sites evaluated, three are publicly held corporations; five are subsidiaries of publicly held corporations; four are private corporations or are sponsored by private entities; three are funded by grants, donations or membership fees; and four are government sponsored through the National Institutes of Health. One Web site, Obesity-online.com, listed no information as to the sponsoring entity.

All of the English-language Web sites except for two (Allhealth.com and Yahoo.com) listed a medical advisory board, editorial board, or committee that has input into and/or oversight of the health content. Some sites list only board member names and credentials, but many of the Web sites also provide detailed biographical information. Allhealth.com does not appear to provide any information regarding a medical or editorial advisory board. But it appears that all of the health content on Allhealth.com is written by outside medical and health experts whose credentials, background, and experience are well documented. Yahoo! also does not indicate that there is a medical advisory board to oversee the site's health content. But according to Yahoo! all of the health information on the site is provided by credible companies in the health care industry.

Table 3.1: Web Sites Included in Evaluation of Quality

English-Language Web Sites	URL Address
Popular General Health	
Allhealth.com	www.allhealth.com
CBS Health Watch	www.cbshealthwatch.com
DrKoop.com	www.drkoop.com
Intelihealth	www.intelihealth.com
Onhealth	www.onhealth.com
WebMD	www.webmd.com
Yahoo!	www.yahoo.com
Condition-Specific	
American Academy of Allergy, Asthma & Immunology	www.aaaai.org
American Cancer Society	www.cancer.org
American Obesity Association	www.obesity.org
Athealth.com	www.athealth.com
Cancernet	www.cancernet.gov
Depression.com	www.depression.com
MyAsthma	www.myasthma.com
National Heart, Lung, and Blood Institute	www.nhlbi.nih.gov
National Institute of Mental Health	www.nimh.nih.gov
National Library of Medicine	www.nlm.nih.gov
Obesity Online	www.obesity-online.com
Oncolink	www.oncolink.com
Spanish-Language Web Sites	
Popular General Health	
Graciasdoctor	www.graciasdoctor.com
Salud	www.salud.com
Salud Latina	www.saludlatina.com
Condition-Specific	
Cancernet	www.cancernet.gov
Centro Peso	www.centropeso.com
National Institutes of Health	www.nih.gov
New York Online Access to Health	www.noah-health.org

Table 3.2: Descriptive Information on English-Language Web Sites Studied

SITE #	WEB SITE	SOURCES OF REVENUE	1999 REVENUES	ADVISORY BOARD	WHO WRITES MEDICAL CONTENT?
1	Allhealth.com (health channel for iVillage.com)	Ivillage, Inc.: publicly held Nasdaq: IVIL Advertising and sponsorship, commissions on sponsor sales, barter transactions	$36.6M	No information	Outside medical experts (listed and vitae posted); featured articles by others indicate authorship
2	Onhealth.com (WebMD Corp.)	WebMD – publicly held Nasdaq: HLTH Transaction services, physician services (including office management products), portal services, other products and services	$102.1M	Medical advisory board (listed and vitae posted)	Experts in various clinical areas (listed and some vitae posted)
3	Intelihealth.com (subsidiary of Aetna U.S. Healthcare)	Aetna, Inc. NYSE: AET InteliHealth is funded by Aetna U.S. Healthcare (core business of Aetna, Inc.) to the extent not funded by revenues from operations Advertising and sponsorship	$26.4B (Aetna U.S. Healthcare net income: $437.3M)	Editorial boards from Harvard Medical School (listed), University of Pennsylvania School of Dental Medicine (listed and vitae posted), review information created by Intelihealth staff	Staff editors and writers (all listed and bios), freelance writers (listed and bios)
4	CBSHealthWatch.com (MedicaLogic/Medscape, Inc.)	MedicaLogic/Medscape, Inc. – publicly held Nasdaq: MSCP Advertising and sponsorship, including banner advertising, grants, and sponsorship of continuing education programs	$11.2M	Editorial board (listed and vitae posted)	Staff writers, physicians and other health care professionals, health writers
5	WebMD.com	WebMD – publicly held Nasdaq: HLTH Transaction services, physician services (including office management products),	$102.1M	Health advisory board (listed and vitae posted)	Medical editors, physicians, nurses, journalists, health communication experts, etc.

Table 3.2: Descriptive Information on English-Language Web Sites Studied

SITE #	WEB SITE	SOURCES OF REVENUE	1999 REVENUES	ADVISORY BOARD	WHO WRITES MEDICAL CONTENT?
		portal services, other products and services			
6	Drkoop.com	Publicly held Nasdaq: KOOP Advertising and sponsorships, content syndication, electronic commerce	$9.4M	Medical advisory board (listed and bios)	Freelance health writers, journalists, physicians, and other health professionals (listed and some vitae posted)
7	Yahoo.com	Publicly held Nasdaq: YHOO Advertisements, business services, barter transactions, electronic commerce transactions	$588.6M	No information	Health information companies: Adam.com, HealthCentral.com, DrKoop.com, Drweil.com, Wellmed, Healthscout, iVillage, eFit, Women.com, Tavolo, INNX Health, Kalive.com, PDR Family Guides, The Little Blue Book
8	Aaaai.org (American Academy of Allergy, Asthma and Immunology)	Unrestricted educational grant from Schering/Key Advertising in Academy News Membership	No information	Public education committee (not listed)	Reviewed by medical editor (not specified)
9	Myasthma.com (Protocol Driven Healthcare, Inc.)	Protocol Driven Healthcare, Inc. Privately held Provides branded interactive self-care Web sites to subscription-based corporate clients Advertising, sponsorship, and links to commercial sites for online shopping	No information	Medical advisory board (listed)	"MyAsthma team" along with leaders in the field (not listed)

Table 3.2: Descriptive Information on English-Language Web Sites Studied

SITE #	WEB SITE	SOURCES OF REVENUE	1999 REVENUES	ADVISORY BOARD	WHO WRITES MEDICAL CONTENT?
10	Nlm.nih.gov	Federal funds	No information	Sites must meet selection guidelines; no specific information on who reviews the sites	Not a source of primary information; links to consumer health information
11	Obesity-online.com	No information	No information	Editorial board (listed)	No information
12	Obesity.org (American Obesity Association)	Nonprofit corporation Donations from individuals, companies, foundations and other sources Membership	No information	Advisory council (listed and vitae posted)	No information
13	Oncolink.com (University of Pennsylvania Cancer Center)	University of Pennsylvania Cancer Center; Aventis Oncology; American Viatical Corp.; Ortho Biotech, Inc.; Varian, Inc.; Janssen Pharmaceutical; grants from NCI and National Action Plan on Breast Cancer; donations	No information	Editorial advisory board (listed)	Physicians, nurses, health writers (listed)
14	Nhlbi.nih.gov	Federal funds	$1.7B	Numerous institute advisory committees/boards	No specific information
15	Bronx-leb.org	Sponsored by Bronx-Lebanon Hospital Center	No information	Medical directors for each specialty (listed)	No information
16	Depression.com (satellite health channel of Planet Rx)	Planet Rx – publicly held Nasdaq: PLRX Prescription drug sales (49% of e-commerce revenue); non-prescription	$9.0M	Medical advisory board (listed and vitae posted)	Team of professional health journalists under supervision of medical

Table 3.2: Descriptive Information on English-Language Web Sites Studied

SITE #	WEB SITE	SOURCES OF REVENUE	1999 REVENUES	ADVISORY BOARD	WHO WRITES MEDICAL CONTENT?
		drugs, personal care products, beauty and spa products, vitamins, herbs and nutrition products and medical supplies Sponsorship			advisory board; medical doctors specializing in depression; in-house pharmacists (some listed and vitae posted)
17	Athealth.com	Privately held Advertising; sponsorship	No information	General advisory board: 6 members – 2 MDs (listed)	Various professionals and organizations; content and articles list authorship, credentials, and often vitae posted
18	Nimh.nih.gov	Federal funds, donations	$854.2M (Total NIMH)	The National Advisory Mental Health Council (NAMHC) advises on activities, policies and programs	NIMH staff writers with MD assistance
19	Cancernet.nih.gov	Federal funds through NCI, donations	$2.9B (Total NCI)	Editorial boards that specialize in cancer information	Experts in oncology and related specialties
20	Cancer.org (American Cancer Society)	Donations, investment income, government grants	$67.2M	Division boards of directors (medical and lay)	Medical editors (MDs), health journalists, oncology nurse specialists

3.3: Descriptive Information on Spanish-Language Web Sites Studied

Site #	Web Site	Sources of Revenue	1999 Revenues	Advisory Board	Who Writes Medical Content?
1	Noah-health.org	Federal Library Services and Technology Act funds granted by New York State Library; individual sponsorship	No information	Scientific advisory board of health professionals (listed and vitae posted)	Librarians and specialists in medical information develop collection of content (listed and vitae posted); author names and/or institutional affiliations are displayed on the linked pages
2	Salud.com	A subsidiary of MillenniumHealth Communications, Inc. Limited information: Transaction services (online marketplace for medical equipment and supplies, online pharmacy), health information offered for syndication	No information	Medical advisory board (listed and vitae posted)	Staff editors and writers: physician and health professionals (listed and vitae posted)
3	Saludlatina.com	No information	No information	No information	No information
4	GraciasDoctor.com	No information	No information	Medical advisory board (listed and vitae posted)	No information
5	Nhlbi.nih.gov	Federal funds	$1.7B*	Numerous institute advisory committees/boards	No specific information
6	Cancernet.nih.gov	Federal funds through NCI, donations	$2.9B**	Editorial boards that specialize in cancer information	Experts in oncology and related specialties
7	Centropeso.com	Created by Productos Roche S.A, manufacturer of pharmaceuticals and diagnostic systems	5.0B CHF (Swiss Franc)+	No information	No information

*Total Budget of NHLBI
** Total NCI
+ Total Roche

With regard to medical content, all of the publicly held Web sites provided information on who writes their health content. The four government sites did not offer specific information about content writers but indicate that content is written by medical experts within the institutes. Three had no information regarding the source of their health content. Of the five remaining sites, three provided information on their health content and two made reference to medical editors or writing teams but gave no specific information.

Spanish-language Web sites. Of the seven unique Spanish-language Web sites evaluated; two are private corporations or are sponsored by private entities; one is funded by grants, donations, and membership fees; two are government sponsored through the National Institutes of Health; and two provide no information as to the nature of their sponsoring entity.

All but two of the Web sites listed a medical advisory board, editorial board, or committee that has input into and/or oversight of the health content. Some sites list only board member names and credentials, but many also provide detailed biographical information. Saludlatina.com and Centropeso.com do not appear to have any information regarding a medical or editorial advisory board. But it appears that all of the health content on Saludlatina.com is written by outside medical institutes, health organizations, or journals, which are documented on the site. Centropeso.com does not indicate the source of any of its health content.

With regard to medical content, two sites (Noah-health.org and Salud.com) out of seven provide information on who writes their health content. The two government sites do not offer specific information about content writers but indicate that content is written by medical experts within the institutes. The remaining three sites (Saludlatina.com, GraciasDoctor.com, and Centropeso.com) have no specific information regarding authorship of their health content. As mentioned, Saludlatina.com cites the majority of its health content from outside sources. GraciasDoctor.com occasionally cites an author but gives no more information than a name. Centropeso.com gives no information regarding the authors of its health content.

Selecting Expert Panelists

To assess the quality of health-related information on the Web site, a series of condition-related topics and corresponding consumer-oriented questions deemed essential for consumers to know were identified. To generate these topic areas and questions for each of the four medical conditions, small panels consisting of clinical experts and representatives from patient advocacy organizations for each of the four conditions were assembled. Each panel consisted of 3-4 experts recruited for their national reputation in condition of interest, clinical or scientific experience, familiarity with national guidelines, and current research. Appendix B lists the expert panelists for each medical condition.

Developing Condition-Related Topics and Consumer Questions

We developed *condition-related topic*s and corresponding *consumer questions* through a structured, multi-step process involving both RAND staff and the expert panels. RAND staff developed an initial list of topic areas and consumer questions based on a review of national guidelines and scientific literature as well as informal discussions with clinical experts in each of the medical fields (American College of Radiology, 1998; Williams and Wilkins, 1996; National Task Force on the Prevention and Treatment of Obesity, 2000; National Heart, Lung and Blood Institute, 1998; Mulrow et al, 1999; Linde et al, 1998; Cole et al, 1999; Mynors et al, 1995; Wilson et al, 1998; American Psychiatric Association, 1994).

Concurrent with the RAND staff process, each expert panel member was asked to submit five to ten consumer-oriented questions that would reflect the concerns of patients, their families, or lay persons seeking information on the study conditions. Panelists were asked to consider questions from each of three categories: (1) condition-related topics about which there is broad expert consensus and for which clear guidelines exist; (2) clinical topics about which there is uncertainty; (3) recent important developments in screening, diagnosis, or treatment of the condition. An example of the first category of question is when to start breast cancer screening using mammography in women over 50. An example of the second category relates to the issue of breast cancer screening using mammography in women age 40-49. An example of the third category of question is the recall of fentermine/phenfluramine for weight loss.

The lists developed by RAND staff and the expert panelists were combined to create the master preliminary list. Panelists were asked to rate each topic area and corresponding question independently on acceptability of the topic or question and level of consensus. Acceptability was rated as: (0) not acceptable to include; (1) acceptable to include: (2) preferable to include; (3) essential to include. Level of consensus was rated as: (0) no consensus; (1) minimal level of consensus; (2) moderate level of consensus; (3) high level of consensus. RAND staff collected the ratings from each panelist and developed a score on acceptability and level of consensus for each topic area and consumer question.

Panelists then met on a series of conference calls to narrow the list of topics and questions through discussion. The list agreed to on the conference call as the top five to ten questions were rated a second time by panelists independently using the acceptability and level of consensus ratings described above. RAND staff compiled the results of the second round of ratings and sent the results to the panelists. Each panel reassembled on a conference call to review the ratings for each of the topics and questions. The final five to seven questions used in this study to assess quality were agreed to by consensus of the panelists during the conference call. Panelists were asked to consider whether the topic or question was relevant to consumers, important for them to have an answer to, and necessary to find on any Web site offering information about the condition. The final set consisted of 26 condition-related topics and 36 consumer-oriented questions across the four medical conditions. They can be found in Appendix C.

Development of Condition-Related Clinical Elements

Based on extensive literature reviews, the panels then developed a series of standardized clinical elements (concepts that should be addressed) for the topics and questions. These concepts were, in effect, the sort of clinical information that consumers should expect to find for a given condition on a Web site. The expert panelists were also involved in developing these elements. RAND staff began by drafting a set of proposed condition-related clinical elements based on a review of national guidelines and the scientific literature. These condition-related clinical elements were then circulated to the expert panelists for comment. RAND staff compiled comments from all of the experts and sent the final set of clinical elements out to the panelists for final approval. For example, for the topic of breast cancer screening, four clinical elements were developed. These included: women older than 50 should have mammograms every 1 to 2 years; early detection of breast cancer improves outcomes; most breast cancers occur in women without a family history; and a lack of consensus exists about the need for or appropriate interval of mammography in women from age 40 to 49 years. A final set of 100 clinical elements were developed. Appendix D lists the condition-related topics and the clinical elements.

Retrieving Health Information from English- and Spanish-Language Web Sites

Between October 18-30, 2000, and November 6-13, 2000, four abstractors (two monolingual in English, two bilingual in English and Spanish) each spent 90 minutes independently reviewing each Web site using efficient DSL connections. The time limit was based on pilot data and studies showing that consumers spend about 30 minutes looking for health information during a given search session (Cyber Dialogue, 2000). Each abstractor was asked to retrieve any information related to the consumer questions developed by the expert panelists. Abstractors did not receive any of the condition-related clinical elements prior to conducting each search. All searches began at a common starting place (for example, condition-specific page or home page if a condition-specific page did not exist). On average, about 65% of pages selected from the Web sites were common among abstractors. The search results were saved using a software application called CatchTheWeb® (Math Strategies, Greensboro, North Carolina). This software application allowed project researchers to accurately save, abstract, and manage Web pages for use at a later date.

Overall, 2,662 Web pages (defined by the programmer's end-of-page mark) containing 21,711 printed pages of material were retrieved from Web sites across the four conditions; 19,529 printed pages were retrieved from the English-language sites and 2,182 printed pages were retrieved from the Spanish-language sites.

All materials from each search were then assembled into separate notebooks with features identifying the source site removed before the review process. Each notebook contained the materials retrieved from a single search and an accompanying standardized rating form. The 78 unique English-language notebooks averaged 250 pages and ranged from 21 to 547 printed pages. The 32 unique Spanish-language notebooks averaged 68 pages and ranged from 8 to 366 printed pages.

Evaluating the Information Retrieved from the Web Sites

A total of 34 (30 monolingual in English, four bilingual in English and Spanish) physicians from around the United States were recruited to review the abstractor-retrieved materials. All physician-reviewers were board-eligible or board-certified in family medicine, general surgery, internal medicine (including allergy and immunology, hematology and oncology, infectious diseases, pulmonary and critical care), or pediatrics. We gave the physician-reviewers the Web site materials in the form of a notebook (described above); each notebook contained the content found by one abstractor on one Web site and an accompanying rating form. Each reviewer rated materials for one to four conditions. Each physician-reviewer rated at least one notebook; no reviewer rated more than five notebooks of material for any condition. No physician-reviewer rated materials from the same site twice. Forty English-language (51%) and 14 Spanish-language (44%) notebooks were selected at random for a second review to evaluate inter-rater reliability.

Rating Form Development. To obtain ratings from the reviewers, RAND staff developed a rating form for each condition. The form listed the topic area, the corresponding consumer question, and the clinical elements. Ratings were obtained at the level of the clinical element on both coverage and accuracy. The rating form also provided space for reviewer to write notes about any conflicting information identified during his/her review of the materials in the notebook.

Rating Coverage. Reviewers were first asked to rate the coverage for each clinical element on a three point scale (0 = not addressed, 1 = minimally addressed, and 2 = more than minimally addressed). "Not addressed" meant there was no reference to the issue on any page of the notebook. For

example, under screening for breast cancer, if no mention was made of the use of mammography for screening, that would be rated as "not addressed." Under therapeutic modalities for childhood asthma, for example, if inhaled corticosteroids or inhaled beta$_2$-agonists were not mentioned, that would be rated as "not addressed." If, under treatments for depression, there was no mention of antidepressant medications, that would be rated as "not addressed." And for obesity, if there was no mention of body mass index, the clinical element related to a definition of obesity would be rated as "not addressed."

"Minimally addressed" meant the concept was mentioned at least briefly. For example, under screening for breast cancer, if mammography was mentioned as a way to identify early breast cancers, but no mention was made of who should have mammograms, how often they should be done, or their utility in reducing breast cancer mortality, this would be considered minimal coverage. For example, if, under triggers that contribute to an exacerbation of childhood asthma, indoor triggers are mentioned, but specific examples such as cockroach antigens, tobacco smoke, and dust mites are not mentioned, this would be minimal coverage. Under treatments for depression, if antidepressants are mentioned, but no mention of side effects related to the anti-depressants is found, this would be minimal coverage. For obesity, if body mass index was mentioned, but the ranges for overweight (25-29.9) and obesity (30+) were not provided, this would be classified as minimal coverage.

More than minimally addressed meant most or all of the elements listed in the topic areas were at least mentioned and the level of explanation was more than cursory. For example, reviewers would rate coverage as more than minimal if a Web site mentioned that screening mammography was the best way for breast cancer to be detected early in women older than 50, or that breast cancer may be detected earlier by mammography than by physical examination, or if a detailed discussion of the pros and cons of mammography and the appropriate ages were provided. For childhood asthma, an example of a more than minimally covered item would be under inhaled medications that reduce inflammation of the airways and the two available types of this kind of medication (inhaled steroids and Cromolyn) are listed, as well as their indicated uses (patients with persistent asthma who are having uncontrolled symptoms). An example of more than minimal coverage under treatments for depression would be that various types of treatments including drug and non-drug therapy as well as side effects were mentioned.

Rating accuracy. For each clinical element that was at least minimally addressed, reviewers also rated the accuracy of content on a three-point scale (0 = content was mostly incorrect, 1 = content was mostly correct, and 2 = content was completely correct). Reviewers were always instructed to give the higher score if they were uncertain about the rating.

Report on conflicting information. After completing ratings of coverage and accuracy, reviewers of English-language sites were asked to list instances of conflicting information found during their review. These conflicts did not necessarily involve the set of clinical elements for which coverage and accuracy were evaluated. All examples of conflicting information were collected from the reviewed Web site materials. Six categories of conflicting information were identified. The categories included: (1) treatments, (2) diagnosis, (3) definitions, (4) adverse effects, (5) etiology and risk factors, and (6) incidence and prevalence. Two project staff physicians then independently rated whether the examples of conflicting information were minor, significant, or potentially dangerous. Examples that were identified as significant or potentially dangerous by both raters were included in

the final analysis. Disagreements were settled by discussion between raters with a tendency to rate the conflict as less significant.

Authorship, Dating, and Currency. Web sites were rated by RAND staff according to the HON (Health on The Net Foundation) code's criteria of authorship (whether authors and their affiliations and credentials were clearly identified) and dating (whether the date the Web site's materials were created or updated was specified). When dates were specified, the currency, defined as the most recent date of modification, was coded.

Units of Analysis and the Derivation of Outcome Variables. The unit of analysis was the rating form for analyses of coverage, accuracy, and conflicting information. For overall (cross-topic) analyses of coverage and accuracy, clinical element-level scores were aggregated at the reviewer (rating form) level into a single observation of each outcome variable per rating form. Summary scores at the condition or Web site level were derived as averages of these reviewer-level observations. For topic-level summaries of coverage and accuracy, clinical element-level of scores were aggregated at the topic level within reviewers, then averaged across reviewers. The unit of analysis was the Web page for the study of authorship, dating, and currency, excluding Web site home pages.

General analytic approach. All analyses were conducted separately for the English- and Spanish-language Web sites. All statistical tests were two-sided, and were assessed for significance at the 0.05 level. Measures were tested for variation by condition and by site within condition. A two-stage test procedure was used to examine variation in each outcome by these independent categorical variables. First, an omnibus or overall test of the association was performed. If the omnibus test confirmed that variation in the outcome of interest was statistically significant for a given categorical variable (condition or site), a series of two-sample follow-up tests was performed. These follow-up tests compared the outcome at each level of the categorical variable to the overall level of the outcome.

The omnibus tests employed were one-way analysis of variance (ANOVA), the Kruskal-Wallis rank-sum test, and the chi-squared test of homogeneity for measures that were normally, ordinally, and nominally distributed, respectively. Two-sample t-tests, Wilcoxon rank-sum tests, and chi-squared tests of homogeneity were the corresponding follow-up tests.

Analyses of coverage. The extent to which Web sites provided information relevant to each of the questions/topics and the condition was calculated as the average proportion of condition-related clinical elements that the reviewers rated as: (1) not covered, (2) minimally covered, (3) more than minimally covered. Quality rating forms contained multiple ratings (corresponding to clinical elements) of coverage and accuracy using the three-point ordinal scales. For purposes of analysis, global measures were computed across clinical elements within a given rating form.

Three global or summary measures of coverage were computed across all ratings of coverage within a rating form: (1) the proportion of condition-related clinical elements that were rated as not covered, (2) the proportion of condition-related clinical elements that were rated as minimally covered, and (3) the proportion of condition-related clinical elements that were rated as more than minimally covered.

Analyses of accuracy. Accuracy was assessed only for items on which coverage was rated as minimally or more than minimally covered. Accuracy was calculated as the average proportion of condition-

related clinical elements rated by reviewers as: (1) mostly incorrect, (2) mostly correct, or (3) completely correct. Accuracy was assessed only for items on which coverage was rated in the highest category. One global measure of accuracy was computed across such items: the proportion of covered items that were rated as completely correct.

Combined score for coverage and accuracy. We computed the proportion of clinical elements for each topic rated by reviewers as both more than minimally covered and completely accurate. A global measure combined coverage and accuracy at the clinical element level, computing the proportion of all condition-related clinical elements that received scores in the top category in both coverage and accuracy. Note that these dependent measures are not binary at the rating-form level but are quasi-continuous variables constrained between 0 and 1.

Simulations of consumer search for sites with extensive topic coverage. To simulate the experience of a consumer trying to identify a health Web site with extensive coverage of a particular health topic, we created a model based on the coverage results of the studied Web sites and applied it to a subset of eight condition-related topics (two per condition) that were thought to be of particular consumer interest. This model assumed that the consumer searched from a large universe of health Web sites until finding a site with more than minimal coverage for 75% of the five to seven clinical elements corresponding to the topic of interest. For the purpose of this simulation, we assumed that the levels of coverage observed among the studied sites were representative of the levels of coverage in the larger universe of health sites that addressed the condition to which the topic of interest applied. We further assumed that the log-odds of the levels of more-than-minimal coverage for the studied sites came from a normally distributed universe of log-odds coverage proportions in the larger universe of sites.[2] Taken together, these assumptions allowed us to estimate the proportion of sites in the universe of sites addressing that condition that would provide such a level of coverage for a selected topic. This in turn allowed us to estimate the average (expected) number of sites that would have to be visited and examined (never returning to the same site) before finding a site with the desired level of topic coverage. When the expected average number of sites needed exceeded the total number of sites studied for a language and condition (nine or ten for English, four for Spanish), we concluded that there was not clear evidence of the existence of any sites within that language that provide extensive coverage of the topic within a single site.

Analyses of conflicting information. For each of the six conflict categories a binary variable was created, indicating the presence or absence of any significant or potentially dangerous conflict of a given category on each rating form. The two-stage testing procedure described above was used to assess variation in the prevalence of conflicting information by condition among English-language Web sites.

Analyses of authorship and dating. Two binary indicators of whether a content-containing Web page listed an author or a date of creation or modification were combined to construct a three-level ordinal scale for Web pages: (1) neither author nor date, (2) either author or date, (3) both author and date. The two-stage testing procedure described above was employed to assess variation in the proportion of Web pages for a given site or condition that had: (1) neither author nor date and (2) both an author and a date.

[2] Because log-odds for 0% and 100% are infinite, we replaced Web site scores of 0% and 100% with 5% and 95%, respectively, for purposes of estimation.

Inter-rater reliability. Two measures of physician-reviewer inter-rater reliability of Web site quality ratings were computed. A standard measure of reliability, the correlation in ratings between physician-reviewers examining identical material retrieved from the same Web site, was calculated. Because we wanted to assess the sensitivity of physician-reviewer ratings to variation in the retrieved material (for example, the material retrieved by abstractor 1 versus abstractor 2 on the same Web site), a second, more stringent, measure was also computed. This was the correlation between ratings of different physician-reviewers examining different notebooks of material from the same Web site, and it represents the inter-rater reliability of the entire process we used to evaluate sites. We computed 16 inter-rater reliabilities by the standard rule and 16 by the stringent rule for each language: one for every combination of the four conditions and the four assessments (any coverage, more than minimal coverage, completely correct, and the combination of more than minimal coverage and complete correctness). Thirty reviews were included in each calculation of inter-rater reliability on Spanish-language Web sites. The standard inter-rater reliability was high: 0.90 or greater in all instances, averaging 0.96 for both English- and Spanish-language sites. The more stringent inter-rater reliability was also high for English Web sites (average reliability of 0.77) and fair for Spanish Web sites (average reliability of 0.60).

Results

We found variation in the quality of information available among conditions, topic areas, and Web sites.

Coverage and Accuracy of Selected Health Topics on English-Language Web Sites

Across English-language Web sites, the clinical elements (identified by expert panelists as important for a Web site to include) that were more than minimally covered varied significantly by condition: 67% for breast cancer, 43% for childhood asthma, 53% for depression, and 40% for obesity (Table 3.4, $p<0.05$). There was also statistically significant variation in coverage among Web sites within conditions. For example, within breast cancer, rates of more than minimal coverage ranged from 31% to 90% (Table 3.4, $p<0.05$). Six of ten breast cancer sites, two of ten depression sites, and no childhood asthma sites or obesity sites, provided more than minimal coverage of two-thirds of the clinical elements.

Accuracy of health information was generally high across English-language Web sites. Among the clinical elements that were at least minimally covered, the average percentage that were completely correct was 91% for breast cancer, 84% for childhood asthma, 75% for depression, and 86% for obesity (data not shown).

We found significant variation among English-language Web sites in the proportion of clinical elements that were both more than minimally covered and completely accurate. For example, the percentage of condition-related clinical elements covered more than minimally and completely correctly for depression ranged from 13% to 73% among English-language Web sites (Table 3.4).

Coverage of Selected Health Topics on English-Language Web Sites by Condition

Breast cancer. Among English-language sites, materials related to three topics (screening, risk assessment, and treatment) were more than minimally covered and completely correct 70% of the time (Table 3.5). Topics not covered most often included alternatives to standard medical and surgical treatments (28%) and evaluation of a palpable breast mass (18%). If the selected topics were addressed, they tended to be accurate most of the time, with ratings ranging from 86% (breast

Table 3.4: Comparison of Coverage and Accuracy of Selected Health Topics on English-Language Web Sites

English Web Sites	Average Percentage (%) of Clinical Elements by Web Site			
	No Coverage	Minimal Coverage	More Than Minimal Coverage	More Than Minimal Coverage and Completely Correct
Breast Cancer (Overall)**	**16**++**	**17**	**67**++**	**63**++**
Oncolink.com	3++	7	90++	85++
Cancernet.nih.gov	6	10	84	80
Webmd.com	4	14	82	80
Cancer.org	10	18	73	73
Drkoop.com	20	12	69	69
Onhealth.com	9	18	74	68
Intelihealth.com	18	22	61	53
Allhealth.com	24	22	55	47
CBSHealthWatch.com	27	25	47	45
Yahoo.com	39--	29	31--	27--
Childhood Asthma (Overall)**	**27**	**30**	**43**--**	**33**--**
NHLBI.nih.gov	20	25	55	52
Myasthma.com	23	30	47	44
Drkoop.com	21	26	53	43
Onhealth.com	16	31	52	42
Intelihealth.com	32	27	41	38
Webmd.com	14	36	50	36
CBSHealthWatch.com	28	28	44	31
Aaai.org	32	35	32	26
Allhealth.com	51--	27	23	22
Depression (Overall)**	**20**--**	**27**	**53**	**44**
NIMH.nih.gov	0	19	81	73++
Intelihealth.com	13	19	68	65
Webmd.com	10	30	60	56
CBSHealthWatch.com	17	27	57	52
Depression.com	18	27	56	48
Drkoop.com	20	33	48	44
Onhealth.com	18	21	62	41
Athealth.com	18	37	46	35
Allhealth.com	23	33	40	13--
Yahoo.com	58	22	17--	13--
Obesity (Overall)**	**35**--**	**25**	**40**--**	**37**--**
CBSHealthWatch.com	21	19	60	59
Intelihealth.com	20	23	57	52
NLM.nih.gov	30	17	53	51
Webmd.com	27	30	43	42
Obesity.org	42	14	43	40
Drkoop.com	20	38	42	39
Onhealth.com	27	42	31	30
Allhealth.com	39	33	28	27

Table 3.4: Comparison of Coverage and Accuracy of Selected Health Topics on English-Language Web Sites

English Web Sites	Average Percentage (%) of Clinical Elements by Web Site			
	No Coverage	Minimal Coverage	More Than Minimal Coverage	More Than Minimal Coverage and Completely Correct
Obesity (Overall) (cont'd)**	**35**[--]	**25**	**40**[--]	**37**[--]
Obesity-online	54[--]	11	34	27
Yahoo.com	70[--]	20	10	7
Overall Average	**25**	**25**	**51**	**45**

**Weighted by number of reviews. Overall scores are an average of unrounded scores, and therefore do not correspond exactly to the average of rounded site scores.

[++] Significantly better performance than condition average ($p<0.05$)

[--] Significantly worse performance than condition average ($p<0.05$)

Table 3.5: Evaluation of Breast Cancer Information on English-Language Web Sites

Breast Cancer		Average Percentage of Selected Condition-Related Topics for Ten English-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct
1. Risk assessment and use of tamoxifen for risk reduction	Are there any medications I can take to reduce my risk of getting breast cancer?	10	12	78	89	73
2. Screening	No one in my family has had breast cancer. Do I still need breast exams and mammograms? When should I start having regular mammograms? Do I need one every year?	10	11	79	86	69
3. Evaluation of a palpable breast mass	I have a lump in my breast. What should be done to check this?	18	25	57	93	57
4. Treatment	If I have Stage I or Stage II breast cancer, which is better treatment: mastectomy or lumpectomy plus radiation? Where can I get information about breast cancer clinical trials?	13	14	73	92	70
5. Alternatives to standard surgical and medical therapies	Which alternative therapies (such as acupuncture, herbs, or homeopathy) can help me fight breast cancer?	28	23	49	96	49

cancer screening) to 96% (alternative therapies). There was statistically significant variation between English-language Web sites (Table 3.4). One site (Oncolink.com) performed statistically better than average.

A consumer randomly searching a large universe of Web sites addressing breast cancer (according to the simulated consumer search for extensive topic coverage described earlier in this chapter) might be expected to find extensive coverage[3] of the topic regarding treatment of breast cancer within two Web sites but would be expected to have to visit four sites before finding extensive coverage of evaluation of a palpable breast mass.

Childhood asthma. There was more variation in coverage among the seven selected topics for childhood asthma. Overall coverage for all condition-related topics was generally lower (Table 3.5). The topics receiving the best coverage related to therapeutic modalities and side effects (65%) and the etiology of asthma (46%). Topics not covered most often included symptoms suggestive of poorly controlled asthma (48%), initial management of severe asthma (33%), and common symptoms (33%). Topics addressed tended to be completely accurate, with ratings ranging from 72% (signs of poorly controlled asthma) to 98% (etiology). Three percent of materials related to the selected topics contained specific factual inaccuracies. For example, one site describes cockroaches as the leading cause of asthma among children.

No topic area received a combined score (more than minimal coverage and completely accurate) of more than 50%. No Web site performed statistically better than the condition average for childhood asthma (Table 3.4).

A consumer randomly searching Web sites addressing childhood asthma would be expected to search four sites before finding extensive coverage of information related to symptoms suggestive of poorly controlled asthma.

Depression. Coverage of the seven condition-related topics ranged widely. Topics related to the etiology of depression were more than minimally covered 97% of the time, whereas topics related to the type of provider to see for depression were more than minimally covered 13% of the time (Table 3.6). Topics related to antidepressant medications were more than minimally covered two-thirds of the time. Accuracy of coverage ranged from 68% (treatment) to 90% (etiology). Three percent of materials related to the selected topics contained specific factual inaccuracies. As an example, one depression site stated that omega-3 fatty acid deficiencies cause major depressive disorders.

Coverage of topics that were both more than minimally covered and completely correct ranged from 8% (who should evaluate depression) to 87% (etiology). One Web site (nimh.nih.gov) performed statistically better than average (Table 3.4).

A consumer randomly searching Web sites addressing depression would be expected to find extensive coverage of information related to antidepressant medications within two searched sites.

*Obesity***.** Among English-language Web sites, topics covered most often related to health risks (59%) and indications for weight loss (48%). Topics not covered most often related to safety and

[3] More than minimal coverage for at least 75% of the indicators for this topic.

Table 3.6: Evaluation of Childhood Asthma Information on English-Language Web Sites

Childhood Asthma		Average Percentage of Selected Condition-Related Topics for Nine English-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Symptoms	What are the common symptoms of asthma in children?	33	26	41	89	36
2. Poorly controlled asthma	I have been told by a doctor that my child has asthma. S/he has difficulty breathing at night and uses an inhaler every day. Does this mean that my child's asthma is not well controlled?	48	29	23	72	18
3. Therapies and adverse effects	What should I do about my child's asthma, especially if it is not well controlled? Are there any medicines or special equipment that my doctor can prescribe? Do they have any side effects?	13	22	65	76	48
4. Initial management of severe asthma	How do I know if my child is having life-threatening symptoms? What should I do?	33	46	21	84	19
5. Risk factors	Could certain exposures in the indoor and/or outdoor environment have caused or made my child's asthma worse? What can be done to identify, eliminate, or diminish factors in the environment that can worsen my child's asthma symptoms?	29	32	39	84	33
6. Etiology	What causes asthma? Is it curable?	32	22	46	98	46
7. Expectations from therapy	What should I expect from my child's asthma treatment?	23	41	36	90	36

effectiveness of dietary supplements (61%) and risks and benefits of popular diets (49%). Accuracy was generally good, ranging from 78% (indications for weight loss) to 96% (safety and effectiveness of certain dietary supplements). One topic (health risks of obesity) was more than minimally covered and accurate more than half of the time (Table 3.7). No Web site statistically performed better than average (Table 3.4).

A consumer randomly searching Web sites addressing obesity would be expected to search seven sites before finding extensive coverage of materials related to weight-reduction surgery.

Coverage and Accuracy of Selected Health Topics on Spanish-Language Web Sites

On Spanish-language sites, more than half of the clinical elements identified by expert panelists as important for a Web site to include were not covered (Table 3.9). The average percentage of clinical elements not covered varied significantly by condition: 49% for breast cancer, 33% for childhood asthma, 61% for depression, and 69% for obesity (Table 3.9). Levels of coverage varied significantly by condition ($p < 0.05$) but not within condition. More than minimal coverage of any of the clinical elements was rare among Spanish-language sites (39% for breast cancer, 27% for childhood asthma, 15% for depression, and 16% for obesity).

Coverage and Accuracy of Selected Health Topics on Spanish-Language Web Sites by Condition

Breast cancer. Among the five condition-related topics selected for evaluation for breast cancer, only two were covered more than minimally more than 50% of the time (screening and evaluation of a palpable breast mass) (Table 3.10). Topics not covered most often related to the treatment options for Stage I and Stage II breast cancer (61%) and alternatives to standard surgical and medical treatments (90%). No Web site performed statistically better than average.

Childhood asthma. Overall coverage of the seven condition-related topics was low (Table 3.11). One condition-related topic achieved more than minimal coverage 40% of the time (symptoms). Accuracy was also more variable; between 38% and 61% of minimally covered topics were scored as completely correct. The topic with the highest level of coverage related to the symptoms of childhood asthma (44%). No Web site performed statistically better than average.

Depression. Lack of coverage on all condition-related topic areas was particularly striking (Table 3.12). Four of the condition-related topics were more than minimally covered less than 10% of the time (antidepressant medications, role of counseling, suicidal ideation, and evaluation). Web sites provided more than minimal and completely correct coverage on what to do if an individual were experiencing suicidal ideations 5% of the time (Table 3.12). No Web site performed statistically better than average.

Obesity. More than minimal coverage on all obesity topic areas ranged from zero to 31% (Table 3.13). Topics covered most often included materials related to definitions and indications for weight loss (31%) and physical activity and prevention (29%). Topics not covered most often included risks and benefits of popular diets (100%) and safety and effectiveness of dietary supplements (100%). Accuracy was variable, ranging from 50% (availability of drugs approved for weight loss) to 81% (definitions and indications for weight loss). One topic (definitions and indications for weight loss) was more than minimally covered and accurate 30% of the time. No Web site performed statistically better than average (Table 3.9).

Table 3.7: Evaluation of Depression Information on English-Language-Web Sites

Depression		Average Percentage of Selected Condition-Related Topics for Ten English-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Symptoms	I've been feeling a little sad lately. How do I know if I'm depressed?	13	15	72	82	61
2. Treatment	What are the most effective treatments for depression?	17	11	72	68	56
3. Antidepressant medications	If my doctor recommends an antidepressant medication for the treatment of my depression, how long should I take it for? What should I expect and when will I start to feel better?	16	17	67	78	55
4. Role of counseling	When should I consider psychological counseling instead of or in addition to medication?	31	29	40	73	33
5. Suicidal ideation	I feel so depressed I've thought about suicide. What should I do?	12	46	42	84	37
6. Evaluation	Whom should I see for evaluation and treatment of my depression? A primary care doctor, a psychiatrist, or a psychologist/ therapist?	33	54	13	71	8
7. Etiology	What causes depression?	0	3	97	90	87

Presence of Conflicting Health Information on English-Language Web Sites

In the course of reading the Web site material, many reviewers noted the presence of conflicting information within a Web site. As mentioned previously, these conflicts were not necessarily related to the set of condition-related topics for which coverage and accuracy were measured. For English sites only, we calculated the proportion of times raters noted at least one significant conflict of condition-related information during their review. For example, one childhood asthma Web site reported in one place that using inhaled steroids does not stunt growth in children, and elsewhere it reported that using inhaled steroids does stunt growth in children.

Overall, just more than half of Web site reviews revealed one or more conflicts of a clinically important nature (Table 3.14). Conflicts most often involved treatment (35%) and diagnosis (13%). Materials on depression most commonly had conflicts on treatment, whereas breast cancer materials most commonly contained conflicts on diagnosis ($p< 0.05$). Appendix E lists the examples of types of conflicts noted by the physician-reviewers.

Authorship, Dating, and Currency of Content on English- and Spanish-Language Web Sites

Approximately 65% of English-language materials listed both an author (institutional, individual, or both) and a date (Table 3.15). Forty-six percent of all English-language materials had been created within the past year, and 45% of those dated materials had been modified within the past one to three years (Table 3.17). Approximately 9% of the materials retrieved from the English-language Web sites contained no evidence of any author or date of publication or modification.

By contrast, 14% of the Spanish-language materials specified both an author (institutional, individual, or both) and a date (Table 3.16). Seventeen percent of all the Spanish-language materials had been created within the past year, and 32% of those dated materials had been modified with the past one to three years (Table 3.17). Approximately 44% of the materials retrieved from the Spanish-language Web sites contained no evidence of any authorship or date of publication or modification.

Discussion

We examined several dimensions of Web site quality (availability of key information, accuracy, identifiable authorship, and currency) related to four common health problems (breast cancer, childhood asthma, depression, and obesity). Although we found thousands of pages of material related to the key clinical topics and questions, we found gaps in the availability of key information.

What Did We Find?

Most sites provided at least minimal coverage of 75% of the condition-related topics we looked for on the sites (Table 3.8). Some sites, however, provided very little information with up to 70% of condition-related topics completely uncovered. Only four of the English-language Web sites (oncolink.com, cancernet.nih.gov, webmd.com, and nimh.nih.gov), and none of the Spanish-language Web sites, provided more than minimal coverage for at least 80% of the condition-related topics. Breast cancer topic areas were covered significantly more often than all other conditions; topic areas about childhood asthma and obesity were covered significantly less often. Even fewer Spanish-language Web sites provided more than minimal coverage of topics with information that was completely correct. Although the accuracy of information presented was fairly high, more than half of the Web sites reviewed revealed one or more conflicts of a clinically important nature, such as about a treatment choice. About 65% of all English-language materials contained an author and a date, and most of the materials were published within one to three years. By contrast, 14% of all

Table 3.8: Evaluation of Obesity Information on English-Language-Web Sites

Obesity		Average Percentage of Selected Condition-Related Topics for Ten English-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Definitions and indications for weight loss	How do I know if I need to lose weight?	36	16	48	78	42
2. Health risks	What are the health risks of being overweight/obese?	12	29	59	90	56
3. Risks and benefits of popular diets	What should I consider before starting on a low-carbohydrate, high-protein, high-fat diet like the Atkins plan?	49	34	17	87	17
4. Physical activity/ prevention	What is the value of physical activity for (a) promoting weight loss, (b) maintaining weight at current levels, and (c) general health?	20	34	46	94	43
5. Medications endorsed for weight loss	Should I consider weight-loss drugs, and if so, what prescription and non-prescription drugs are currently available?	47	25	28	83	25
6. Surgery	Who should consider weight-loss surgery, what are the risks, and how well does it work?	32	22	46	85	44
7. Safety and effectiveness of dietary supplements	Can herbal supplements containing ephedrine plus caffeine help me to safely lose weight?	61	19	20	96	19

Table 3.9: Coverage and Accuracy of Condition-Related Clinical Elements on Spanish-Language Web Sites [#]

Spanish-Language Web Sites	Average Percentage of Clinical Elements by Web Site			
	No Coverage %	Minimal Coverage %	More Than Minimal Coverage %	More Than Minimal Coverage and Completely Correct %
Breast Cancer (Overall[**])	**49**	**12**	**39[++]**	**39[++]**
Cancernet.nih.gov	21	3	76	76
Salud.com	46	10	44	43
Saludlatina.com	68	12	21	21
Graciasdr.com	61	20	20	20
Childhood Asthma (Overall[**])	**33[++]**	**40**	**27**	**23**
NIH.gov	33	34	34	30
Salud.com	28	45	27	26
Saludlatina.com	41	41	19	16
Graciasdr.com	33	44	23	13
Depression (Overall[**])	**60**	**25**	**15[--]**	**15**
Salud.com	63	10	26	24
Noah.edu	54	30	17	15
Graciasdr.com	58	29	11	5
Saludlatina.com	73	19	7	2
Obesity (Overall[**])	**69[--]**	**15**	**14**	**15**
Salud.com	53	25	22	22
Saludlatina.com	80	3	17	17
Graciasdr.com	68	14	18	15
Centropeso.com	73	18	10	10
Overall Average	**53**	**23**	**24**	**22**

[#] Differences between sites within condition were not statistically significant.

**Weighted by number of reviews. Overall scores are an average of un-rounded scores, and therefore do not correspond exactly to the average of rounded site scores.

[++] Significantly better performance than condition average ($p<0.05$)

[--] Significantly worse performance than condition average ($p<0.05$)

Table 3.10: Evaluation of Breast Cancer Information on Spanish-Language Web Sites

Breast Cancer		Average Percentage of Selected Condition-Related Topics for Four Spanish-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Risk assessment and use of tamoxifen for risk reduction	Are there any medications I can take to reduce my risk of getting breast cancer?	45	14	41	83	36
2. Screening	No one in my family has had breast cancer. Do I still need breast exams and mammograms? When should I start having regular mammograms? Do I need one every year?	27	12	61	100	61
3. Evaluation of a palpable breast mass	I have a lump in my breast. What should be done to check this?	30	20	50	97	50
4. Treatment	If I have Stage I or Stage II breast cancer, which is better treatment: mastectomy or lumpectomy plus radiation? Where can I get information about breast cancer clinical trials?	61	12	27	88	27
5. Alternatives to standard surgical and medical treatments	Which alternative therapies (such as acupuncture, herbs, or homeopathy) can help me fight breast cancer?	90	0	10	100	10

Table 3.11: Evaluation of Childhood Asthma Information on Spanish-Language Web Sites

Childhood Asthma		Average Percentage of Selected Condition-Related Topics for Four Spanish-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Symptoms	What are the common symptoms of asthma in children?	25	31	44	61	40
2. Poorly controlled asthma	I have been told by a doctor that my child has asthma. S/he has difficulty breathing at night and uses an inhaler every day. Does this mean that my child's asthma is not well controlled?	42	33	25	57	25
3. Therapies and adverse effects	What should I do about my child's asthma, especially if it is not well controlled? Are there any medicines or special equipment that my doctor can prescribe? Do they have any side effects?	27	44	29	56	27
4. Initial management of severe asthma	How do I know if my child is having life-threatening symptoms? What should I do?	40	50	10	38	10
5. Risk factors	Could certain exposures in the indoor and/or outdoor environment have caused or made my child's asthma worse? What can be done to identify, eliminate, or diminish factors in the environment that can worsen my child's asthma symptoms?	32	42	26	47	21
6. Etiology	What causes asthma? Is it curable?	31	36	33	52	23
7. Expectations from therapy	What should I expect from my child's asthma treatment?	44	39	17	56	14

Table 3.12: Evaluation of Depression Information on Spanish-Language Web Sites

Depression		Average Percentage of Selected Condition-Related Topics for Four Spanish-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Symptoms	I've been feeling a little sad lately. How do I know if I'm depressed?	45	37	18	64	14
2. Treatment	What are the most effective treatments for depression?	42	22	36	55	24
3. Antidepressant medications	If my doctor recommends an antidepressant medication for the treatment of my depression, how long should I take it for? What should I expect and when will I start to feel better?	77	14	9	94	9
4. Role of counseling	When should I consider psychological counseling instead of or in addition to medication?	64	27	9	53	9
5. Suicidal ideation	I feel so depressed I've thought about suicide. What should I do?	55	40	5	50	5
6. Evaluation	Who should I see for evaluation and treatment of my depression? A primary care doctor, a psychiatrist, or a psychologist/ therapist?	84	16	0	33	0
7. Etiology	What causes depression?	27	9	64	75	45

Table 3.13: Evaluation of Obesity Information on Spanish-Language Web Sites

Obesity		Average Percentage of Selected Condition-Related Topics for Four Spanish-Language Web Sites				
		Coverage			Accuracy	Combined
Condition-Related Topic	Corresponding Consumer Question	No Coverage, %	Minimal Coverage, %	More Than Minimal Coverage, %	Completely Correct, %	More Than Minimal Coverage and Completely Correct, %
1. Definition and indications for weight loss	How do I know if I need to lose weight?	50	19	31	81	31
2. Health risks	What are the health risks of being overweight/obese?	23	50	27	52	21
3. Risks and benefits of popular diets	What should I consider before starting on a low-carbohydrate, high-protein, high-fat diet like the Atkins plan?	100	0	0	-	0
4. Physical activity/ prevention	What is the value of physical activity for (a) promoting weight loss, (b) maintaining weight at current levels, and (c) general health?	54	17	29	63	29
5. Medications endorsed for weight loss	Should I consider weight-loss drugs, and if so, what prescription and non-prescription drugs are currently available?	79	11	10	50	8
6. Surgery	Who should consider weight-loss surgery, what are the risks, and how well does it work?	90	6	4	71	4
7. Safety and effectiveness of dietary supplements	Can herbal supplements containing ephedrine plus caffeine help me to safely lose weight?	100	10	0	-	0

Table 3.14: Presence of Conflicting Information on English-Language Web Sites

	Proportion of Reviews Noting Any Conflict (%)	Proportion of Reviews with Any Conflicting Information by Category (%)					
		Treatment	Diagnosis	Definition	Side Effects	Etiology and Risk Factors	Incidence and Prevalence
All Medical Conditions	**53%**	**35%**	**13%**	**7%**	**5%**	**5%**	**4%**
Breast Cancer	43%	20%-	27%+	0%	0%	3%	3%
Childhood Asthma	52%	45%	0%-	3%	10%	7%	3%
Depression	73%	50%+	17%	7%	7%	10%	10%
Obesity	43%	27%	7%	17%	3%	0%	0%

+Significantly more reviews for this condition noted conflicts of this type than for the mean condition ($p < .05$)
-Significantly fewer reviews for this condition noted conflicts of this type than for the mean condition ($p < .05$)

Table 3.15: Evidence of Authorship and Currency of English-Language Web Sites

English-Language Web Sites	Percent of Web Pages with Neither Author nor Date	Percent of Web Pages with Either Author or Date*	Percent of Web Pages with Both Author and Date
Breast Cancer (Overall)**	**10**	**18**	**72**
Oncolink.com	0++	5	95++
Onhealth.com	2++	3	95++
Webmd.com	2++	3	95++
Cancernet.nih.gov	0++	18	82
CBSHealthWatch.com	2	19	79
Cancer.org	2	21	77
Drkoop.com	8	17	75
Intelihealth.com	9	47	44--
Allhealth.com	6	70	24--
Yahoo.com	89--	3	8--
Childhood Asthma (Overall)**	**6**	**37**	**57**
Onhealth.com	2	3	95++
Webmd.com	8	0	92++
CBSHealthWatch.com	0	22	78++
Myasthma.com	2	20	78++
Drkoop.com	11	23	66
Aaaai.org	0	61	39--
NHLBI.nih.gov	29--	38	33--
Allhealth.com	5	63	32--
Intelihealth.com	8	91	1--
Depression (Overall)**	**8**	**26**	**66**
NIMH.com	0++	7	93++
Depression.com	3	8	89++
Webmd.com	6	9	85++
Drkoop.com	1++	17	82++
Onhealth.com	6	15	79++
CBSHealthWatch.com	4	29	67
Intelihealth.com	1++	62	37--
Allhealth.com	2	74	24--
Athealth.com	36	40	24--
Yahoo.com	96--	4	0--
Obesity (Overall)**	**12**	**25**	**63**
Webmd.com	3++	7	90++
CBSHealthWatch.com	3++	10	87++
Obesity.org	0++	17	83++
Onhealth.com	1++	16	83++
NLM.nih.gov	17	11	72
Drkoop.com	10	19	71
Intelihealth.com	3++	46	51--
Allhealth.com	3++	82	15--
Obesity-online	55--	32	13--

Yahoo.com	100--	0	0--
Overall Average	**9**	**26**	**65**

** Weighted by number of pages retrieved. Overall scores are an average of unrounded scores, and therefore do not correspond exactly to the average of the rounded site scores.
* Contains evidence of either an author or a date but not both.
++ Significantly better performance than the condition average ($p<0.05$)
-- Significantly worse performance than the condition average ($p<0.05$)

Table 3.16: Evidence of Authorship and Currency on Spanish-Language Web Sites

Spanish Web Sites	Percent of Web Pages with Neither Author nor Date	Percent of Web Pages with Either Author or Date*	Percent of Web Pages with Both Author and Date
Breast Cancer (Overall)**	**30**	**38**	**32**
Cancernet.nih.gov	0++	0	100++
Salud.com	0++	76	24
Saludlatina.com	0	100	0--
Graciasdr.com	86--	14	0--
Childhood Asthma (Overall)**	**52**	**48**	**0**
Salud.com	25++	75	0
NIH.gov	33	67	0
Saludlatina.com	67	33	0
Graciasdr.com	76--	24	0
Depression (Overall)**	**33**	**38**	**29**
Noah-health.org	0++	0	100++
Saludlatina.com	6++	35	59++
Salud.com	26	72	2-
Graciasdr.com	79--	21	0--
Obesity (Overall)**	**59**	**40**	**1**
Salud.com	13++	83	4
Saludlatina.com	50	50	0
Graciasdr.com	78--	22	0
Centropeso.com	92--	8	0
Overall Average	**44**	**42**	**14**

** Weighted by number of pages retrieved. Overall scores are an average of unrounded scores, and therefore do not correspond exactly to the average of the rounded site scores.
++ Significantly better performance than the condition average ($p<0.05$)
-- Significantly worse performance than the condition average ($p<0.05$)

Table 3.17: Currency of Dated Material on English- and Spanish-Language Sites

Web Sites	Percentage of Dated Web Pages
English-Language	
Content created or updated in past year	46%
Content most recently updated 1-3 years ago	45%
Content most recently updated 3-5 years ago	7%
Content not updated within past 5 years	2%
Spanish-Language	
Content created or updated in past year	17%
Content most recently updated 1-3 years ago	32%
Content most recently updated 3-5 years ago	51%
Content not updated within past 5 years	0%

Spanish-language materials contained an author and a date, and just half of those materials were published within one to three years.

Findings from this study suggest that consumers using the Internet may have a difficult time finding information on a health problem. Some of the gaps were particularly striking. For example, less than half of the Spanish-language materials explained that mastectomy and lumpectomy plus radiation are equivalent treatments for early-stage breast cancer. If people rely on the Internet to help guide their health decisions, these deficiencies in information could have serious consequences.

Can We Believe What We Found?

Critics of this study might ask whether the questions we used are really of interest to consumers, whether the selected Web sites we evaluated are representative of the material available, whether the answers we used to judge comprehensiveness and accuracy were reasonable, whether physician reviewers had access to all of the available information on a site, and whether their assessments were reliable. Here we discuss each of these points in more detail.

To standardize the assessment of content on Web sites, we used a group of experts to identify the key questions consumers should have been able to answer when seeking information on a particular topic. The experts included both health providers who treat patients with these conditions and consumer advocates who represent the interests of patients with these problems. We have provided the questions so that readers of this report can judge for themselves whether the questions are relevant to consumers. A survey of consumers with the conditions would have been extremely useful but was beyond the scope of this study. Because this study was not a natural experiment (for example, using consumers to search for information and testing their knowledge after such a search), we cannot draw conclusions about what people actually encounter when they search for information, or how well they are able to interpret the information they find.

For the most part, general health sites were selected because of their widespread popularity among consumers. The condition-specific sites, by contrast, are generally less frequently used. These sites were selected largely because they represented a different type of site than some of the most popular ones. Within each language and condition, we compared the average performance of condition-specific sites to the average performance of general health sites with respect to combined coverage and accuracy (the proportion of clinical elements covered more than minimally and with complete accuracy). For breast cancer, a clear pattern was apparent. Within English-language sites, three of the four highest average scores were obtained by the three cancer-specific sites; in Spanish, the best score was obtained by the one cancer-specific site. The performance of the selected set of both English- and Spanish-language breast cancer-specific Web site(s) was significantly better than the performance of the selected set of general health Web sites within the corresponding language on the topic of breast cancer ($p < 0.05$ in both cases). No such pattern was apparent or statistically significant among the remaining three conditions. This may be because for the other conditions, even the best-performing Web sites offered only modest coverage of topics.

Perhaps we found poor performance because our condition-related clinical elements to the questions were too demanding. Some might represent a standard that is "too high," and we cannot be sure that a different research group, assisted by different panelists, would not have generated different concepts. But our clinical panelists were instructed to take the perspective of patients, not physicians, in determining what information ought to be available. Panelists were also instructed to

avoid concepts that were arcane, and we also avoided controversial concepts, except when assessing whether uncertainty or controversy was properly communicated.

Another potential criticism is that because we abstracted relevant information from each site and presented it for review in printed, hard-copy form, physician reviewers did not have access to all relevant materials available on the site. We abstracted information from the sites both to make the review task manageable and to make the Web site being reviewed anonymous (so that reviewers were neither positively nor negatively influenced by knowledge of the source). While we used trained searchers to gather material for the reviewers, they collected what could fairly be described as a sample rather than the entire universe of material available on each Web site. But the searchers were well trained and were given more time than most consumers spend looking for specific information. If our abstractors could not uncover the material in 90 minutes, it is unlikely that the average consumer could do so.

Finally, although we provided the reviewers with the clinical elements, critics might posit that the assessments were largely subjective. To evaluate the level of agreement among reviewers, a random sample of half of the sites was reviewed by a second physician. We found very high levels of inter-rater reliability, suggesting that the assessments were comparable.

Where Does This Leave Us?

Given the substantial variation in coverage of key topics across Web sites, consumers should probably not rely on a single Web site to answer all of their condition-specific questions.

Consumers should not assume that even well-designed and comprehensive-appearing Web sites contain all essential information on a health topic. There may be gaps. If consumers are truly interested in finding comprehensive answers to their questions, they may need to devote more time to the search than is commonly the case and they must be willing to sort through many different sites—perhaps as many as ten different Web sites to find all relevant information.

Conflicts are not uncommon, and consumers are at risk for becoming confused or misinformed. Much of the conflict probably results from the methods by which information is updated—adding new information without systematically reviewing existing text to remove conflicting information. Conflicts can probably never be eliminated—it would be harder to do so on the Web than with standard multi-authored textbooks because of the multi-dimensional layering of electronic information. Therefore, the Web probably should not serve as the final arbiter of health care information for consumers—they need access to a professional who can clarify inconsistencies and reconcile conflicts.

4. Readability Assessment of Health Information on the Internet

Purpose

The Internet has the potential to reduce barriers in access to information for patients but only if the online material can be read and understood by many types of individuals. Most studies estimate that more than half of U.S. residents read at the ninth-grade level or lower (Davis, 2000; Root and Stableford, 1999). Health-related information has been shown to be more difficult to comprehend than most other types of information (Root, et al., 1999).

The reading ability of patients varies widely and is generally lower than the level of school they have completed. One study of English-speaking diabetic patients found that while 60% could understand information written at the sixth-grade level, only 21% could understand information written at the ninth-grade level (Overland, et al., 1993). Other studies have found a median reading level of ninth to tenth grade in emergency department patients (Williams, et al., 1996) and a median reading level of seventh to eighth grade in cancer patients (Foltz and Sullivan, 1996), patients in urban clinics (Wilson, 1995), and parents of pediatric patients at a university hospital (Davis, et al., 1994). In one study of hospitalized patients, only 7% could comprehend information written at the fifth-grade level, and just 30% could comprehend material written at the ninth-grade level (Estey, et al., 1994).

Limited reading skills may be more prevalent among certain patient and population sub-groups. For example, according to the 1993 National Adult Literacy Survey, 75% of welfare recipients (including but not limited to Medicaid beneficiaries) read at or below the eighth-grade level and 50% read at or below the fifth-grade level (Kirsch, et al., 1993). Immigrants and refugees from less developed countries may be even more likely than their U.S.-born counterparts to have low educational attainment and, as a result, limited reading skills. Among recent Central American immigrants and refugees from El Salvador and Guatemala, only slightly more than 20% reported having completed high school (Lopez, 1996). Among foreign-born Hispanics living in greater Los Angeles, 10% reported no schooling, 3% reported elementary school attendance, 21% reported at least some high school (but no college), and only 5% reported a college degree (Cheng and Yang, 1996).

In this study, we assess the readability of written information from 19 English- and seven Spanish-language health sites. Specifically, we wanted to know: What grade-level reading ability is required to understand health information regarding four common medical conditions on English- and Spanish-language health Web sites?

Methods

Assessing Readability

There are several methods of assessing the readability of a document. The most direct way to measure readability is to administer a comprehension test based on the document of interest to a group of readers of known reading ability. Readability can also be measured by the judgment of a literacy expert. A third approach uses reading formulas, rather than experts or test subjects. Reading formulas are mathematical equations that estimate the reading level of a document based on the words used and the length of sentences.

The methods using test subjects and experts are more costly and time-consuming but also more precise than reading-level formula methods (Klare, 1974). The reading-level formulas can be thought of as automatic approximations of the other methods. In this study, time and resource constraints dictated the use of formula-based methods.

Readability Formulas Employed

We conducted a literature search to identify reading formulas appropriate for Spanish- and English-language documents. Although we found references to numerous readability formulas, few were appropriate for both English and Spanish documents.

Three readability assessment methods were applied to the text from the Spanish and English Web sites: the Fry Readability Graph (FRG), the SMOG grading formula, and the newer Lexile Framework®. The first and third methods are applicable to both English and Spanish documents; only the third is currently implemented in software.

The FRG has been validated for Spanish- and English-language documents (Gilliam, et al., 1980; Fry, 1969, 1977). The FRG uses three sample passages of text, each exactly 100 words in length, from the beginning, middle, and end of the source document. The grade level is computed as a function of the number of sentences and words contained in the three samples of text. Application of the FRG to Spanish-language documents is similar to its application to English-language documents, with the exception of syllable counting. In Spanish an adjustment compensates for the fact that Spanish text contains more syllables per word than English text of the same reading level. (Gilliam, Peña, and Moutain, 1980).

Unlike the FRG, the SMOG grading formula is applicable only to English-language documents.

The SMOG uses three passages of ten sentences each from the beginning, middle, and end of the source. The reading level is a function only of the number of polysyllabic words (words with three or more syllables) in the sampled text, with more polysyllabic words corresponding to higher reading levels[4]. The SMOG grading formula has been used widely and has been adopted by the National

[4] The reading level is estimated by the formula $\mathrm{SMOG} = 3 + \sqrt{\mathrm{PSC}}$, where PSC is the average polysyllable count per ten sentences.

Cancer Institute as the preferred method for assessing the readability of patient communications after a comprehensive review of advantages and disadvantages of alternative readability formulas (Romano, 1979).

The Lexile Framework® is a relatively new software program that estimates the readability level of a document based on two factors: average sentence length and word familiarity[5]. Passages consisting of shorter sentences are assumed to be easier to read than passages consisting of longer sentences. Passages consisting of familiar (commonly used) words are assumed to be easier to read than passages consisting of unfamiliar words (Wright and Strener, 1998). Word familiarity is measured by the frequency with which a given word is used in written United States school texts of various grade levels (Carroll, et al., 1971). In this study, the Lexile Framework® software was applied to three 10-sentence sample passages drawn from the beginning, middle, and end of the source documents.

Selection of Abstracted Web Site Material

As noted in Chapter 3, some Web sites were searched and abstracted for more than one condition, and all Web sites were abstracted by two different searchers, resulting in multiple abstraction documents for a given site. For each site, a single abstraction document was randomly selected among all available documents for readability analysis (see Table 3.1).

Results

Readability of English-Language Web Sites

For the English-language Web sites, the mean FRG reading grade level was 13.2 (SD=2.1), ranging from 10 to 17 (Table 4.1). The mean SMOG reading grade level for English-language Web sites was 13.6 (SD=0.9), and ranged from 12 to 15. The mean Lexile Framework reading grade level was 11.7 (SD=1.0), and ranged from 10 to 14. Among English-language Web sites, the correlation was 0.61 ($p<0.05$) between the FRG and SMOG grading formula, 0.54 ($p<0.05$) between the FRG and the Lexile Framework, and 0.32 ($p=0.18$) between the SMOG grading formula and the Lexile Framework.

Readability of Spanish-Language Web Sites

For the Spanish-language Web sites, the mean FRG reading grade level was 9.9 (SD=2.5) and ranged from 7 to 13 (Table 4.1). The mean Lexile Framework reading grade level was 10.0 (SD=2.6) and ranged from 6 to 13. The correlation between the FRG and the Lexile Framework among Spanish Web sites was 0.49.

The mean reading grade level for the English-language Web sites was higher (more difficult) than for Spanish-language Web sites, as measured by the FRG ($p<0.05$) (Table 4.1).

[5] The average sentence length and average word frequency are combined to obtain a Lexile scale score using the following formula: $Lexiles = 582 + 1768 * \log(\bar{S}) - 386 * \log(\bar{W})$, where $\bar{S}$ is the log mean sentence length and $\bar{W}$ is the log mean word frequency. The Lexile score is then translated to a grade-level reading difficulty.

Table 4.1: Readability Levels for Selected Web Sites by Language*

Reading Level	English-Language Sites % (n)	Spanish-Language Sites % (n)
Elementary School (Grade Levels 1-8)	0	14 (1)
High School (Grade Levels 9-12_	37 (7)	43 (3)
College (Grade Levels 13-16)	53 (10)	43 (3)
Graduate School (Grade Levels 17+)	11 (2)	0

* Readability levels were determined using the Fry Readability Graph method.

Discussion

This analysis shows that much of the health information available on the Internet is beyond the comprehension of many consumers. All of the English Web site documents assessed had material that required at least a ninth-grade reading level, and more than half presented material at the college.

level.[6] Four of seven Spanish-language sites presented at the ninth-grade reading level or higher[7]. Studies of (English-speaking) patients in various clinical settings suggest a ninth-grade reading level is too high for most patients. The U.S. Department of Health and Human Services recommends that patient education materials not exceed a sixth-grade reading level (U.S. Department of Health and Human Services, 1999). The mismatch between the reading ability of patients and the readability of health-related information on the Internet suggests that for it to become a more effective medium for patient education, the readability of the materials on the Internet must be improved.

This is the first study to examine the reading level of Spanish-language health-related information on the Internet. This aspect of the study has special significance because Spanish-speaking patients face greater barriers to traditional sources of health information than English-speaking patients do (Ginzberg, 1991; Mayberry and Mili, Ofili, 2000). Surveys indicate the number of Spanish-speaking persons currently accessing the Internet for health information is increasing. Further efforts to reduce racial/ethnic disparities in access to the Internet (for example, The Digital Divide) through strategies such as Community Access Centers will probably bring Internet access to greater numbers of Spanish speakers in the near future (U.S. Department of Commerce, 1999).

Limitations of Readability Assessments by Readability Formulas

It is widely acknowledged that reading is an interactive process that occurs between the text and the reader. In fact, research shows that readers use experiences, knowledge, and information-processing skills to comprehend text (Johnston, 1983).

Readability formulas, being strictly text-based, do not address the interactive nature of the reading process. Most reading formulas, including those used in this study, employ syntactic and semantic factors and do not directly address factors related to communicating meaning. For instance, readability formulas do not distinguish between written discourse and nonsensical combinations of

[6] Using the FRG.
[7] Using the FRG.

words (Dreyer, 1984). Moreover, formulas cannot assess other critical factors such as the reader's interest, experience, knowledge, or motivation, all of which may influence the reader's ability to comprehend the cognitive task asked by a survey (Duffy, 1985). Other factors related to readability and not assessed by a readability formula include typographical and temporal factors (for example, time allotted to complete the reading task), the cultural appropriateness of materials to intended racial/ethnic and linguistic minority groups, and factors related to the unique nature of the Internet.

Based on the findings of this report and recent research on the reading ability of patients, one thing is clear: There is much work needed to provide English- and Spanish-speaking patients with health-related information on the Internet that is accessible. Currently, the reading level of health-related information provided on the Internet is too high for most English- and Spanish-speaking patients.

5. Conclusions and Recommendations

Use of search engines that lead to health-related Web sites and visits directly to health-related Web sites are two of the most common ways that consumers find health information on the Internet. In this study, we examined what consumers are likely to find when they use these methods to search for health-related information. Specifically we asked:

- How easy is it to find relevant information on important health conditions using search engines?
- What type of information do search engines turn up, and to what extent does this vary by search engine?
- With regard to health-related Web sites, how comprehensive, accurate, and readable is the information they provide?
- Is health-related information as readily available in Spanish as it is in English?

Here we review major study findings, draw conclusions from them, and make recommendations for various parties concerned with health information on the Internet, including consumers and their advocates, health care providers, government policymakers and regulators, and health information providers, such as search engines and Web sites.

Conclusions

Search Engines Perform Differently: Where You Start Matters

One of our findings is that search engines are not interchangeable. The results of a search will vary markedly depending on which search engine is used. Among English-language search engines, we found that the average amount of overlap in the top ten Web sites identified by a structured search with different search engines was just 11% overall; in other words, searches with two different search engines would produce on average of only one common Web site in the top ten listed. This variation results from differences in the way search engines identify relevant material and differences in the methods they use to rank sites.

With the simple search strategies used in this study, search engines were only moderately efficient in locating relevant information on a particular health topic. About one in five links identified by English-language search engines and one in eight links identified by Spanish-language search engines led to a Web page with content related to the health conditions. More than half of consumers who use the Internet report that they spend about a half-hour looking for health information, so efficiency is an important aspect of performance.

Coverage of Important Health Topics Is Spotty

Coverage of topics varies markedly from Web site to Web site and from topic to topic. Even well-designed sites that appear comprehensive often have important gaps. For example, Web sites that provide information on the causes of depression may provide little or no information on the side effects of antidepressant medications.

Information Presented Is Generally Accurate

Most sites provide content that expert reviewers rated as mostly or completely correct. Indeed, there appeared to be a much greater chance that covered topics will be addressed accurately than there is that important topics will be covered at all. But reviewers did find one or more instances of conflicting information judged to be clinically important in more than half of their reviews. Consumers who carefully read most or all of the information on a site might well be confused when they encounter such conflicting information.

Information on the Internet Is Commercialized

A substantial proportion of the information that Internet users are likely to find on Web sites is promotional—i.e., it sells products or services but is not clearly labeled as an advertisement. The commercial nature of this material is not always clear, so it may be difficult for consumers to weigh whether or not commercial interests potentially bias the information they encounter.

Health Information on the Internet Is Written at a High Reading Level

The reading grade level of most Web-based health information is high. Material in English is generally written at a college reading level, and will therefore be largely incomprehensible to the many consumers who read at the sixth- to ninth-grade levels, and will be difficult even for those who read at tenth- to twelfth- grade levels. One Spanish-language Web site presented health materials at the elementary -school level; all others required at least a ninth-grade reading level.

Coverage of Spanish Language Health Information Is Sparse and of Poor Quality

The deficiency of important health information on Spanish-language Web sites was particularly striking. Coverage is much less extensive than it is on English-language sites. More than half of the selected condition-related topics were not addressed.

Recommendations

To Consumers Using the Internet

1. The Internet health space is unimaginably vast, and navigating through that space is more challenging than you might think. Although search engines provide the prospect of quick access to information, they are only moderately efficient at identifying relevant content. Therefore, it is important to set aside adequate time for a search, and plan to visit several sites. Our searchers found that even finding

material for a single consumer question took 10 to 15 minutes per Web site with a high-speed Internet connection. As many as four to six sites must be visited to adequately address many questions.

2. Be aware that sites will not necessarily provide a comprehensive picture of what you need to know about a condition. These sites can only supplement consultation with a health care professional. But they may provide you with information that helps you ask good questions and understand much better what your doctor tells you.

3. Do not be surprised if you find information that is conflicting or difficult to understand. Seek out a health care provider to help you interpret what you find.

To Consumer Advocacy Groups

1. Advocacy groups could help consumers by "adopting" one or two relevant sites and periodically screen content for coverage and accuracy. One could possibly imagine commercial arrangements among e-health companies (which would produce the content) and advocacy organizations (which would arrange for impartial review by experts).

2. Advocacy groups could also help consumers by pressing for improvements in site content and presentation to make information more complete, accurate, and accessible, and by referring consumers to better sites.

To Health Care Providers

1. Be aware that patients are exposed to a vast and often confusing array of information, which can sometimes be helpful and sometimes a hindrance to the process of providing good care.

2. Professional societies could work on ways to organize physicians, pharmacists, or a whole new group of professionals to provide a more formal interpretive function. In the current health business environment, this may require attention to the mechanisms by which such services could be reimbursed.

3. Specialty societies could provide key clinical content to be placed on Web sites. Much in the way specialty societies have become involved in developing and promulgating guidelines, their involvement in writing and approving clinical content on the Web could significantly improve the coverage, accuracy, and presentation of the material. Specialty societies should work with advocacy organizations to ensure that the material they develop addresses important patient concerns and facilitates patient decision making on when to seek medical care.

To Internet Health Information Providers

1. Web site content providers could commission clinical panels of experts to review coverage, accuracy, and factual conflicts before material is put before the public. They could also work with consumer advocacy organizations to make sure that frequently asked consumer questions are addressed and language used is lay-friendly.

2. Provide information at lower reading grade levels on consumer-oriented, health-related Web sites. Currently, the U.S. Department of Health and Human Services recommends that patient education materials not exceed a sixth-grade reading level. This standard should be recommended to English- and Spanish-language providers of health-related information on the Internet.

3. Provide for systematic review of clinical content by experts, and incorporate a requirement for such review into standards for quality assessments of health-related Internet sites. There are several major private sector efforts to improve the quality of health sites: Health on the Net (HON), Hi-Ethics, eHealth Ethics Initiative, and the AMA Guidelines. Our results support the need for such review, which should be undertaken under the auspices of an independent party that is unrelated to the Internet health information provider and that does not provide such information itself.

4. Incorporate readability standards into overall quality assessments of health-related Web sites. Readability standards are not currently included in quality assessments such as the HON code. Because readability is an integral part of providing accessible health-related information on the Internet, explicit standards should be articulated. Methods of assessing readability should be disclosed. Readability assessments should be made regularly and disclosed to consumers choosing among health Web sites. Assessments should be made on Spanish- and English-language Web sites and the results should be published and disseminated in English and Spanish.

To Policymakers and Regulators

1. The best English-language Web sites are far better than the best Spanish-language sites. Major gains in quality of Spanish-language information (and very likely other non-English-language information) can be achieved by translating and culturally adapting what is now available in English.

2. Continue to fund high-quality sites, especially for consumer audiences that are not well served by free-market forces alone.

3. Put greater effort into publicizing and increasing access to high-quality government sites.

4. Fund research on effective communication of health information in Web-based format to readers with a wide range of reading levels.

A key challenge across all of these recommendations is the extent to which the market for health information will reward those who provide the highest-quality material. This is parallel to the problems faced in the health care delivery system—the focus is frequently on cost rather than quality, or on ease of access rather than likelihood of a good outcome. As the business environment for the Internet evolves, it will serve the public well if mechanisms can be found to sort the higher-quality from the lower-quality information providers. It seems unlikely that market forces alone will be sufficient. The solutions may have to come from cooperation among foundations and the not-for-profit sector, including consumer advocacy organizations. Ultimately, we will all benefit from the easy availability of better health information.

Appendices

Appendix A: List of Related Medical Terms for the Search Engine Study

Breast Cancer	
English-Language	**Spanish-Language**
American Cancer Society (ACS)	American Cancer Society (ACS)
Biopsy	biopsia
BRCA (genetic susceptibility to cancer)	BRCA (gen de susceptibilidad al cáncer)
Breast	mama
Cancer	cáncer
Chemotherapy	quimioterapia
Clinical trial	prueba clínica
Estrogen postive	estrógeno positivo
Estrogen negative	estrógeno negativo
Fine needle aspiration	aspiración mediante una aguja fina
Herceptin (brand name)	Herceptin (brand name)
Hormone therapy	hormonoterapia
Implants	implantes subpectorales
Lump (Lumpectomy)	tumor o tumorectomía (extirpación del tumor)
Lymphedema	limfedema
Malignancy	
Mammogram	mamograma
Mammography	mamografía
Mastectomy	mastectomía
Mass	masa
National Cancer Institute (NCI)	Instituto Nacional del Cáncer (NCI)
National Institutes of Health (NIH)	Institutos Nacionales de la Salud (NIH)
Neoplasm	
Oncology	oncología
Raloxifene	
Self-examination	autoexploración
Sentinel node	
Support group	grupo de apoyo
Surgery	cirugía
Tamoxifen (brand name)	Tamoxifen (brand name)
Taxol (brand name)	Taxol (brand name)
Tumor	tumor
Ultrasound	ultrasonido
Childhood Asthma	
English-Language	**Spanish-Language**
Aerochamber (brand name)	Aerochamber (brand name)
Air pollution	contaminación del aire
Albuterol	Albuterol
Allergic [adj]	alérgico(a) [adj]
Allergy, Allergies	alergia, alergias
Allergic rhinitis	rinitis alérgica
American Lung Association	American Lung Association
Asthma	asma
Asthmatic	asmático(a) [adj.]
Asthma attack	ataque de asma
Asthma self-management, education	educación de asma

Asthmacort	
Atrovent (brand name)	Atrovent (brand name)
Beclomethasone	beclometasona
Beclovent	Beclovent (brand name)
Bronchodilator	broncodilatador
Carpet Cleaning	limpieza de alfombras
Cockroaches	cucarachas
Cromolyn sodium (aka: sodium cromoglycate)	cromoglicato de sodio
Dust mites	garrapatas de polvo
Exercise-induced symptoms	síntomas de asma inducida por ejercicio
Flovent (brand name)	Flovent (Brand name)
Inhaler	inhalador
Intal (brand name)	Intal (brand name)
Insprease	
Ipratropium bromide	ipratropium bromide
Leukotriene inhibitors	inhibidores leucotrienos
MDI (metered dose inhaler)	Inhalador de Dosis Medida
National Heart, Lung, and Blood Institute (NHLBI)	National Heart, Lung, and Blood Institute
Nebulizer	nebulizador
Night symptoms	síntomas nocturnos
Peak flow	flujo "pico"
Pediapred (brand name)	Pediapred (brand name)
Prednisolone	prednisolona
Prednisone	prednisona
Proventil (brand name)	Proventil (brand name)
Serevent (brand name)	Serevent (brand name)
Singulaire (brand name)	Singulaire (brand name)
Smoking	fumar
Spacer	espaciador
Steroids (includes: inhaled and oral)	esteroides (incluye: inhalados y orales)
Theophylline	teofilina
Vanceril (brand name)	Vanceril (brand name)
Ventolin (brand name)	Ventolin (brand name)
Wheezing	espiración forzada y rápida
Depression	
English-Language	**Spanish-Language**
Antidepressant	antidepresivo
Anxiety	ansiedad
Bipolar	bipolar
Blues	depresión
Breakdown	ataque de nervios
Counseling	terapia
Depression	depresión
Dysthymia	distimia
ECT therapy	terapia electroconvulsiva
Elavil (brand name)	Elavil (brand name)
Insomnia	insomnio
Manic depression	depresión maniaca
Mental health	salud mental
Mood disorder	trastorno del estado emocional
MAO (monoamine oxidase) inhibitor	inhibidores de la monoaminooxidasa (IMAO)
Nervous [adj.]	nervioso(a) [adj.]
Nortriptyline (brand name)	Nortriptyline (brand name)

Panic attacks	ataques de pánico
Paxil (brand name)	Paxil (brand name)
Prozac (brand name)	Prozac (brand name)
Psychiatric disorder	trastorno psiquiátrico
Sadness	tristeza
SAME: S-adenosyl-methionine	
Seasonal affective disorder (SAD)	
Serotonin	serotonina
SSRI (selective serotonin reuptake inhibitors)	
St. John's wort	
Stress	estrés
Suicide	suicidio
Zoloft (brand name)	Zoloft

Obesity

English-Language	**Spanish-Language**
Abdominoplasty	abdominoplasti
Appetite supressant	inhibidor de apetito
Anorexia	anorexia nerviosa
Binge [v.]	hartarse de algo [v.]
Bulimia	bulimia
Body mass index (BMI)	índice de masa corporal
Body image	imagen del cuerpo
Dexatrim (brand name)	Dexatrim (brand name)
Diet	dieta
Eating disorder	
Exercise	ejercicio
Fat	tejido adiposo
Fen-phen (brand name)	Fen-Phen (brand name)
Gastric Bypass Surgery or Gastroplasty	cirugía gástrica
Jenny Craig	Jenny Craig
Liposuction	liposucción
Lipoplasty	lipoplasti
Meridia (brand name)	Meridia (brand name)
Nutrition	nutrición
Nutrisystem (brand name)	Nutrisystem (brand name)
Obese [adj.] or obesity [n.]	obeso(a) [adj.], obesidad [n.]
Orlistat (brand name)	Orlistat (brand name)
Overweight	obesidad
Purge	purgar
Sleep apnea	apnea en estados de acidosis y de vasconstricción de las arteriales pulmonares en el sueño
Slim	delgado(a) [adj.]
Surgery	cirugía
Thin	delgado(a) [adj.]
Tummy tuck	cirugía plástica abdominal
Weight or weight loss	peso
Xenical (brand name)	Xenical (brand name)

Appendix B: Expert Panelists

Breast Cancer

Sara Collina, J.D.
Senior Policy Analyst
National Breast Cancer Coalition
1707 L Street NW
Washington, DC 20036

Patricia Ganz, M.D.
Professor, Schools of Medicine and Public Health
Director, Division of Cancer Prevention and Control Research
Jonsson Comprehensive Cancer Center
650 Charles Young Drive South
Box 956900
Room A2-125 CHS
Los Angeles, CA 90095-6900

Eric Winer, M.D.
Dana Farber Cancer Institute
44 Binney St.
Boston, MA 02115

Childhood Asthma

Noreen Clark, M.D.
Dean
Marshall H. Becker Professor of Public Health
University of Michigan
School of Public Health
109 S. Observatory Street
Ann Arbor, MI 48109-2020

Lani Wheeler, M.D.
Pediatric and School Consultant
Anne Arundel County Department of Health
3 Harry S Truman Parkway
Annapolis, MD 21401-7085

Depression

Joshua Freedman, M.D.
President, Psychiatric Society for Informatics
Assistant Clinical Professor of UCLA Department of Psychiatry and Biobehavioral Sciences
921 Westwood Blvd., Suite 220
Los Angeles, CA 90024

Appendix B: Expert Panelists

Depression (cont'd.)

Michael Gitlin, M.D.
UCLA Department of Psychiatry and Biobehavioral Sciences
10920 Wilshire Blvd., Suite 1070
Los Angeles, CA 90024

Daniel Mendelson
Program Director
National Depressive and Manic-Depressive Association
730 N. Franklin St., Suite 501
Chicago, IL 60610

Obesity

Richard Atkinson, M.D.
Professor of Medicine and Nutritional Sciences
Director, Beers-Murphy Clinical Nutrition Center
University of Wisconsin
1415 Linden Dr.
Madison, WI 53706

Steve Phinney, M.D., Ph.D.
Vice President for Clinical Nutrition
Galileo Laboratories, Inc.
5301 Patrick Henry Drive
Santa Clara, CA 95054

David Williamson, Ph.D.
Senior Biomedical Research Scientist
Division of Diabetes Translation (K-68)
CDC
4770 Buford Hwy, NE
Atlanta, GA 30341-2717

Appendix C: Condition-Related Topics, Questions, and Clinical Elements Evaluated on English- and Spanish-Language Web Sites

Condition-Related Topic Area	Corresponding Consumer-Oriented Questions	Number of Clinical Elements Used for Assessing Coverage and Accuracy per Topic Area
Breast Cancer		
1. Risk assessment	Are there any medications I can take to reduce my risk of getting breast cancer?	2
2. Screening	No one in my family has had breast cancer. Do I still need breast exams and mammograms? When should I start having regular mammograms? Do I need one every year?	4
3. Evaluation of a palpable breast mass	I have a lump I my breast. What should be done to check this?	4
4. Treatment	If I have Stage I or Stage II breast cancer, which is better treatment: mastectomy or lumpectomy plus radiation? Where can I get information about breast cancer clinical trials?	4
5. Alternative therapies	Which alternative therapies (such as acupuncture, herbs, or homeopathy) can help me fight breast cancer?	3
	Total Number of Condition-Related Clinical Elements for This Condition	*17*
Depression		
1. Symptoms	I've been feeling a little sad lately. How do I know if I'm depressed?	4
2. Treatment	What are the most effective treatments for depression?	3
3. Antidepressant medications	If my doctor recommends an antidepressant medication for the treatment of my depression, how long should I take it for? What should I expect and when will I start to feel better?	3
4. Role of counseling	When should I consider psychological counseling instead of or in addition to medication?	3
5. Suicidal ideation	I feel so depressed I've thought about suicide. What should I do?	2
6. Evaluation	Whom should I see for evaluation and treatment of my depression? A primary care doctor, a psychiatrist, or a psychologist/therapist?	4
7. Etiology	What causes depression?	1
	Total Number of Condition-Related Clinical Elements for This Condition	*21*

Appendix C: Condition-Related Topics, Questions, and Clinical Elements Evaluated on English- and Spanish-Language Web Sites

Condition-Related Topic Area	Corresponding Consumer-Oriented Questions	Number of Clinical Elements Used for Assessing Coverage and Accuracy per Topic Area
Childhood Asthma		
1. Symptoms	What are the common symptoms of asthma in children?	4
2. Uncontrolled asthma	I have been told by a doctor that my child has asthma. S/he has difficulty breathing at night and uses an inhaler every day. Does this mean that my child's asthma is not well controlled?	2
3. Therapeutic modalities and side effects	What should I do about my child's asthma, especially if it is not well controlled? Are there any medicines or special equipment that my doctor can prescribe? Do they have any side effects?	8
4. Symptoms	How do I know if my child is having life-threatening symptoms? What should I do?	4
5. Risk factors	Could certain exposures in the indoor and/or outdoor environment have caused or made my child's asthma worse? What can be done to identify, eliminate, or diminish factors in the environment that can worsen my child's asthma symptoms?	7
6. Etiology	What causes asthma? Is it curable?	4
7. Expectations from therapy	What should I expect from my child's asthma treatment?	3
	Total Number of Condition-Related Clinical Elements for This Condition	*32*
Obesity		
1. Indications for weight loss	How do I know if I need to lose weight?	6
2. Risks	What are the health risks of being overweight/obese?	4
3. Therapies	What should I consider before starting on a low-carbohydrate, high-protein, high-fat diet like the Atkins plan?	3
4. Prevention	What is the value of physical activity for (a) promoting weight loss, (b) maintaining weight at current levels, and (c) general health?	4
5. Availability of drugs approved for weight loss	Should I consider weight-loss drugs, and if so, what prescription and non-prescription drugs are currently available?	4
6. Treatment	Who should consider weight-loss surgery, what are the risks, and how well does it work?	6
7. Safety and effectiveness of dietary supplements containing ephedra plus caffeine	Can herbal supplements containing ephedrine plus caffeine help me to safely lose weight?	3
	Total Number of Condition-Related Clinical Elements for This Condition	*30*

Appendix D: Clinical Elements by Condition

Breast Cancer	
Condition-Related Topic	**Condition-Related Clinical Elements**
1. Assessment of Breast Cancer Risk and Use of Tamoxifen for Risk Reduction	Risk factors for breast cancer include a family or personal history of breast cancer, early menarche, pregnancy history, and a history of breast biopsies.
	In the short run (meaning up to five years), tamoxifen reduces breast cancer risk in high-risk women.
2. Screening for Breast Cancer	Women over 50 should have mammograms every one to two years.
	Early detection of breast cancer improves outcomes.
	Most breast cancers occur in women without a family history.
	There is a lack of consensus about the need for or appropriate interval of mammography in women from ages 40-49.
3. Evaluation of a Palpable Breast Mass	New breast lumps should be brought to the attention of a physician.
	Mammography and ultrasound are useful in evaluating lumps.
	A negative mammogram does not eliminate the need for further evaluation.
	A persistent, non-cystic (non-fluid-filled) breast mass felt by a physician should be biopsied.
4. Treatment, Including Primary Treatment and Availability of Clinical Trials for Treatment of Advanced Cancers	Mastectomy and lumpectomy plus radiation are equivalent treatments for early stage breast cancer.
	Patient preferences should be considered in treatment decisions around mastectomy versus lumpectomy plus radiation.
	Breast reconstruction is available for women who have mastectomy.
	Clinical trials are available for women with advanced cancer. Some information about finding clinical trials is given.
5. Alternatives to Standard Medical and Surgical Treatments for Breast Cancer	Alternative therapies to treat breast cancer have generally not been subjected to rigorous scientific studies.
	Alternative therapies should not be used as a substitute for proven effective treatments.
	Your physician should be informed of any alternative treatments you are using, including herbs, supplements, and over-the-counter medications.
Childhood Asthma	
Condition-Related Topic	**Condition-Related Clinical Elements**
1. Symptoms of Pediatric Asthma	A child with asthma can experience the following symptoms: (1) cough, (2) wheezing, (3) chest tightness, (4) shortness of breath or difficulty breathing, or (5) an "asthma attack" (pronounced or prolonged presence of these symptoms). (Please note: a Web site that does not include the most important and noticeable symptoms (#2, 4, 5) should not be rated as "more than minimally addressed.")
	These symptoms can be worse at night; triggered by exercise, environmental irritants, changes in weather, or viral illness; or can occur spontaneously at rest.
	Children with asthma can have intermittent symptoms (twice a week or less) or persistent symptoms (more than twice a week).
	Children with intermittent symptoms may have a severe exacerbation.
2. Symptoms Suggestive of Uncontrolled Pediatric Asthma	Children with intermittent symptoms (day symptoms twice a week or less and/or night symptoms twice a month or less) are considered "controlled."
	Children with persistent symptoms (day symptoms and/or need to use a rescue medication more than twice a week or waking up with symptoms during the night more than twice a month) are "not controlled."

Appendix D: Clinical Elements by Condition

3. Pediatric Asthma-Therapeutic Modalities and Associated Side Effects	Bronchodilator medications (e.g., albuterol, Proventil, Ventolin) open the airways (breathing passages.) They are used as "quick relief" or "rescue" medications for patients whose symptoms are intermittent (as defined in Topic Area 1.)
	Inhaled corticosteroids (e.g., beclomethasone, flunisolide, triamcinolone, Beclovent, Vanceril, Flovent, Azmacort) and cromolyn (Intal) are two kinds of inhaled medications that reduce inflammation in the airways. They are used as long-term treatments for patients whose symptoms are persistent or uncontrolled.
	A spacer device will improve delivery of inhaled medications to the lungs. Such devices are required for young children and are strongly recommended for older children and adolescents.
	Peak flow monitoring is a useful way for patients to recognize early signs of worsening asthma.
	Oral steroids are effective for short-term exacerbations but have significant side effects over the long term.
	Inhaled steroids, taken in usual doses, do not affect children's growth. Uncontrolled asthma can retard a child's growth.
	Alternative therapies for asthma (e.g., herbal remedies and chiropractic manipulation) have not been shown to be effective.
	Anti-leukotrienes are a new class drug that might be useful as an add-on to inhaled steroids or to prevent exercise symptoms in children over age 6. The safety and efficacy of these drugs in children under 6 has not been demonstrated.
4. Initial Management of Severe Pediatric Asthma	Some children can die of asthma, especially if the early warning signs of a severe asthma attack are missed.
	Signs of a life-threatening asthma episode include: (1) very difficult breathing, (2) shortness of breath at rest, (3) uncontrolled coughing, (4) severe chest tightness, (5) blueness around the lips or nails, (6) difficulty talking, (7) extreme tiredness, fatigue or (8) unresponsiveness. (To warrant a score of two under coverage, the material must include mention of symptoms # 1, 4, 5, 8).
	Immediate home care for a severe asthma attack includes prompt administration of the child's quick-relief or rescue medication.
	If symptoms of a severe asthma attack are not relieved within ten minutes or if the child's symptoms worsen, the caretaker should call 911.
5. Pediatric Asthma – Risk Factors	Certain indoor allergens and irritants (e.g., tobacco smoke, dust mites, cockroach allergens, cat hair) have been shown to cause worsening of acute asthma in children who are sensitive to these factors but not to cause asthma per se.
	Although pollution is not a proven cause of asthma, persons with asthma can experience more asthma exacerbations on high-pollution days.
	Other indoor allergens or irritants such as mold, animal dander other than cat, pollen, strong odors, etc. have been reported to be associated with worsening asthma symptoms. But there is scientific uncertainty about the role of these factors. (For full credit, must mention both the potential role of these allergens/irritants and uncertainty about their importance.)
	Allergens or irritants that trigger a child's asthma can usually be identified through a careful medical history. Blood and skin tests conducted by an allergy specialist can also be helpful.
	Most children being considered for immunotherapy should be evaluated and followed by an allergy specialist.
	Allergy immunotherapy for children with asthma should only be considered when: (1) there is clear evidence of a relationship between symptoms and exposure to an unavoidable allergen to which the child is sensitive, (2) symptoms occur all year or during a major portion of the year, and (3) the symptoms are not controlled with medications. (All three required for full credit.)

Appendix D: Clinical Elements by Condition

5. Pediatric Asthma – Risk Factors (continued)	Families that are sensitive to tobacco smoke, dust mites, cockroach antigens, or cats should undertake vigorous exposure reduction strategies. (Site must suggest any of the specific exposure reduction strategies listed for three out of four of these irritants for full credit.)
6. Etiology and Risk Factors	The cause of asthma is not known.
	Most experts speculate that asthma may be caused by a combination of genetic (hereditary) and environmental factors (exposures).
	Asthma is not contagious and is not caused by psychological or psychiatric disturbances.
	Although asthma medications can control symptoms, asthma is not curable, given current science.
7. Pediatric Asthma – Expectations from Therapy	In 80-85% of cases, children with asthma can be symptom-free if they follow a preventive medication regimen and/or avoid allergens or irritants to which they are sensitive.
	Even if 100% freedom from symptoms is not possible, the disease can be controlled so that the child experiences minimal symptoms during the day and night.
	Children with asthma should be able to participate in normal activities (school, play, etc.) and parents should not have to lose work or sleep time because of children's asthma symptoms.
Depression	
Condition-Related Topic	**Condition-Related Clinical Elements**
1. Symptoms of Depression	The primary symptoms of depression are persistent low mood, loss of interest and enjoyment, and reduced energy lasting at least two weeks.
	Other symptoms of depression include significant weight, sleep, and appetite changes; anxiety; feelings of worthlessness or inappropriate guilt; diminished ability to think or concentrate or indecisiveness; recurrent thoughts of death or suicidal ideation; apathy; or irritability. A person may have a depressive disorder without having all of these symptoms.
	In older patients (defined as 65 years or older), depression may not always present with low mood as seen in younger patients. Instead, patients may seem apathetic and uninterested in normal activities. Anxiety and memory impairment may also be the principal presenting symptoms.
	Depression should not be regarded as a normal part of aging.
2. Treatments for Depression	Effective treatments for depression include prescription antidepressant drugs, specific psychological treatments (cognitive therapy, cognitive behavioral therapy, and interpersonal therapy), combination therapy, and electroconvulsive therapy (ECT.)
	No antidepressant is superior to another in efficacy or time to response. The choice of medication is based upon side-effect profile or prior response.
	St. John's wort (hypericum perforatum) may be an effective treatment for mild depression. But because of reported drug interactions, patients who are taking other prescription medicines should consult with a physician before starting this preparation.
3. Antidepressant Medications	Antidepressant medications typically begin to work within several weeks. But many patients do not experience substantial benefits for 4-6 weeks, and it may take 3-4 months before people on antidepressants feel completely better.
	Patients with a single episode of acute depression who experience initial improvement should continue to take the medication, usually for 6-12 months after they feel completely better to keep feeling well.
	With antidepressant medicines, many people have some side effects early in treatment (in the first four to six weeks). Most side effects get better in the first month. For some people, the side effects can be bad enough to stop the medicine. Common side effects include anxiety, sexual dysfunction, sleepiness, trouble sleeping, weight gain/loss, restlessness, and nausea.

Appendix D: Clinical Elements by Condition

4. Role of Counseling	For mild to moderate depression, prescription antidepressant drugs and specific psychological therapies are equally effective.
	For moderate to severe depression, prescription antidepressant drugs are more effective than psychological therapies.
	For severe depression, the combination of drug therapy with psychological treatment is probably more effective than psychological therapy alone.
5. Suicidal Ideation	People who have suicidal thoughts but are confident they will not carry out suicide should obtain a medical or psychiatric evaluation promptly.
	People with suicidal thoughts who think there is any chance they might attempt suicide should seek emergency evaluation and help from their physician or at an emergency room.
6. Professional Evaluation of Depression	The best person to see for the evaluation of and treatment for depression is uncertain. To date, no definitive scientific studies have proven which one is best.
	For mild depression, both the initial evaluation and subsequent treatment can be provided by a primary care doctor, psychiatrist, or psychologist/therapist.
	For moderate depression, an individual should see either a primary care physician or a psychiatrist for an initial evaluation; subsequent treatment may be provided by either the evaluating physician or a psychologist/therapist.
7. Etiology - Depression	The causes of depression are uncertain but probably result from a combination of genetic predisposition and childhood and current psychosocial adversity.
Obesity	
Condition-Related Topic	**Condition-Related Clinical Elements**
1. Indications for Weight Loss, Definitions of Overweight and Obesity	Body mass index (BMI, weight in kilograms squared/height in meters squared) is a useful way to determine whether someone is overweight or obese.
	There is a distinction between overweight and obesity; overweight is currently defined as BMI between 25 and 29.9; obesity is defined as BMI>=30.
	Growing evidence suggests that these thresholds may be too high for certain non-Caucasian populations (e.g., Chinese, Japanese, Hispanics).
	Waist circumference is by itself predictive of future morbidity; high-risk cutoffs are 35 inches for women and 40 inches for men.
	The health risks of obesity also depend upon disease conditions (e.g., CAD, DM), cardiovascular risk factors (e.g., family history, LDL cholesterol, hypertension), and other obesity-associated diseases and risk factors (e.g., gallstones, degenerative joint disease).
	Treatment is indicated when the patient meets criteria for obesity, or when the patient meets criteria for overweight and the patient has: (1) established cardiovascular disease or diabetes; (2) >=2 other risk factors, including hypertension, dyslipidemia, smoking, family history of heart disease, age>=45 for men or 55 for women; or (3) a high waist circumference (>35 inches for women or >40 inches for men).
2. Health Risks of Being Overweight and Obese	There is an increase in mortality as BMI exceeds 25; the risk increases rapidly above a BMI of 30.
	Important morbidities associated with obesity include diabetes mellitus, hypertension, abnormal blood lipids, coronary artery disease, and sleep apnea.
	Other morbidities include gastro-esophageal reflux disease, gallstones, urinary stress incontinence, and osteoarthritis.
	In addition to medical morbidities, overweight/obesity can produce limitations in mobility, reduced functional status, and lower overall quality of life.
3. Risks and Benefits of Low Carbohydrate, High Protein Diets	The benefits of Atkins-type diets include: (a) good short-term weight loss, (b) less hunger than a standard low-fat diet, and (c) better short-term control of insulin-resistant states, including type II diabetes and hypertension.
	Initial rapid weight loss is mostly water loss.
	The long-term safety (beyond 6 months) of these diets has not been established.

Appendix D: Clinical Elements by Condition

4. Value of Physical Activity for Weight Loss, Maintenance, and General Health	Regular physical activity results in modest weight loss, especially when combined with a low-calorie diet.
	Physical activity is more effective at maintaining current weight than at reducing weight.
	Physical activity benefits general health and fitness independent of weight loss.
	Physical activity benefits some obesity-related problems (e.g., diabetes and hypertension), independent of weight loss.
5. Availability of Drugs Approved for Weight Loss	Weight-loss drugs are an FDA-approved option for patients with a BMI >=27 (with concomitant risk factors) or >=30 (without risk factors).
	FDA-approved prescription drugs for weight loss include sibutramine (Meridia), orlistat (Xenical), and phentermine (Fastin).
	Phenylpropanolamine (Dexatrim, Acutrim) is an OTC weight loss agent approved for short-term use (<=3 months).
	Phenylpropanolamine (Dexatrim, Acutrim) has been associated with strokes (although the magnitude of the stroke risk is not established).
6. Indications, Risks and Benefits of Weight Loss Surgery	Weight-loss surgery should be considered when the BMI is 40 or higher, or when it is 35-39.9 in the presence of medical co-morbidities (e.g., diabetes, known cardiovascular disease, severe degenerative joint disease, hypertension, GERD, sleep apnea).
	Gastric restrictive procedures (e.g., vertical-banded gastroplasty), gastric bypass, and malabsorptive procedures (e.g., biliopancreatic diversion) have been shown to be effective.
	Patients can achieve substantial weight loss, often 100 pounds or more.
	Gastric bypass is somewhat more effective (in terms of weight loss) than vertical-banded gastroplasty.
	Death and major complication rates following surgery are approximately equal for gastric bypass and vertical-banded gastroplasty. Operative mortality is less than 0.5%, morbidity is approximately 5%, incisional hernia rate is approximately 5%, and small-bowel obstruction occurs in 2% of cases.
7. Safety and Effectiveness of Dietary Supplements Containing Ephedra plus Caffeine	Ephedrine (ephedra) plus caffeine has been shown to be effective as a weight loss supplement.
	Several safety concerns remain, especially for patients with co-morbid conditions that might be worsened by sympathomimetic effects.
	Patients who have heart disease or hypertension should consult a physician before taking this combination.

Appendix E: Examples of Conflicting Information

Breast Cancer	**Conflicting Information**
Significant Conflict	1. Recommends mammography every 2 years, ages 40-50 2. Recommends mammography every year over age 40
Potentially Harmful Conflict	1. Recommends investigating a lump if it doesn't change 2. Urges seeing a provider if a lump is found at all
Childhood Asthma	**Conflicting Information**
Significant Conflict	1. States that inhaled steroids do not stunt growth 2. States that inhaled steroids do stunt growth
Potentially Harmful Conflict	1. Lists oral steroids as a quick-relief medication 2. Lists oral steroids as a long-term medication
Depression	**Conflicting Information**
Significant Conflict	1. St. John's wort is as effective as prescription medications 2. No evidence that St. John's wort is as effective
Potentially Harmful Conflict	1. Serious interactions between tricyclics and SSRIs 2. Prozac mixed with tricyclic is not harmful
Obesity	**Conflicting Information**
Significant Conflict	1. Xenical increases the absorption of fat-soluble vitamins 2. Xenical decreases the absorption of fat-soluble vitamins
Potentially Harmful Conflict	None

Glossary

Condition-related topic: A principal subject area of relevance to patients, their families, or laypersons seeking information on the conditions examined in this study (for example, breast cancer screening).

Clinical element: A component or concept of particular significance within a developed condition-related topic. Clinical elements were developed based on evidence-based guidelines and materials from selected literature reviews. (For example, for the topic "breast cancer screening," four clinical elements were developed. These included: women over 50 should have mammograms every one to two years, early detection of breast cancer improves outcomes, most breast cancers occur in women without a family history, and the lack of consensus about the need for or appropriate interval of mammography in women from age 40-49.)

Inter-rater reliability: The correlation in scores that results when two different reviewers evaluate the same case. It's a measure of agreement between reviewers. A score of 1 indicates perfect agreement between reviewers. A score of 0 indicates no agreement between reviewers.

Link: A hypertext connection that allows an individual to move to another Web page or another part of the same Web page. Links appear as highlighted text or pictures. To follow a link, you click the highlighted material.

Search engine: A program used to search for information on the World Wide Web. Search engines employ various search algorithms to retrieve Web pages with the information sought.

URL (Uniform Resource Locator): A uniform way of naming network resources that allows for pages on the World Wide Web to be linked.

Web page: A document accessible on the World Wide Web.

References

1. Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB. Annual deaths attributable to obesity in the United States. *JAMA* 1999;282:1530-1538.

2. Ambre J, Guard R, Perveiler FM, Renner J, Rippen H. Criteria for Assessing the Quality of Health Information on the Internet. May 4, 1999. Online (http:// hitiweb.mitretek.org/docs/criteria.html). Accessed August 11, 2000.

3. American College of Radiology. Standards for diagnosis and management of invasive breast carcinoma. *CA Cancer J Clin* 1998; 48(2): 83-107.

4. American Medical Association, Council on Scientific Affairs, Ad Hoc Committee on Health Literacy. Health Literacy: Report of the Council on Scientific Affairs. JAMA. 1999;281(6):552-557.

5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders, Fourth ed. Washington, DC: American Psychiatric Association; 1994.

6. Armstrong K, Schwartz J, Asch D. Direct Sale of Sidenafil (Viagra) to Consumers on the Internet. NEJM. 1999;341:1389-1391.

7. Baur C, Deering MJ. Proposed Frameworks to Improve the Quality of Health Website Review. Med Gen Med. September 26, 2000: E35.

8. Beredjiklian PK, Bozentka DJ, Steinberg DR, Bernstein J. Evaluating the source and content of orthopaedic information on the Internet: The case of carpal tunnel syndrome. J Bone Joint Surg Am 2000;82-A:1540-1543.

9. Biermann JS, Golladay GJ, Greenfield MLVH, Baker LH. Evaluation of Cancer Information on the Internet. Cancer. 2000; 86(3): 381-390.

10. Carletta J. Assessing agreement on classification tasks: the kappa statistic. Computational Linguistics. 1996; 22(2): 249-254.

11. Carroll JB, Davies P, Richmond, B. The Word Frequency Book, Boston: Houghton Mifflin; 1971.

12. Cheng L, Yang PQ. The "Model Minority" deconstructed. In R. Waldinger and M. Bozorgmehr, eds. Ethnic Los Angeles (pp.305-344). New York, NY: Russell Sage Foundation; 1996.

13. Cole MG, Bellavance F, Mansour A. Prognosis of depression in elderly community and primary care populations: a systematic review and meta-analysis. Am J Psychiatry 1999; 156:1182-1189.

14. Cyber Dialogue. Cybercitizen Health. Online (http://www.cyberdialogue.com). Accessed January 5, 2001.

15. Davis TC, Mayeaux EJ, Fredrickson D, et al. Reading ability of parents compared with reading level of pediatric patient education materials. Pediatrics. 1994 Mar; 93(3):460-8.

16. Davis TC, Williams MV, Branch WT, Green KW. Explaining Illness to Patients with Limited Literacy. In: Whaley BB, (eds.) Explaining Illness: Research, Theory and Strategies. New Jersey: Lawrence Erlbaum Associates; 2000.

17. Davison K. The quality of dietary information on the world wide web. Clin Perform Qual Health Care 1997;5:64-66.

18. Dreyer G. Readability and responsibility. Journal of Reading. 1984; 27:334-339.

19. Duffy T. Readability formulas: What's the use? In: Duffy T, Walker R, eds. Designing Usable Texts (pp. 113-143). Orlando, FL: Academic Press, Inc.; 1985.

20. Estey A, Musseau A, Keehn L. Patient's understanding of health information: a multihospital comparison. Patient Educ Couns. 1994 Aug; 24(1):73-8.

21. Eysenbach G, Yihune G, Lampe K, Cross P, Brickley D. Quality Management, Certification and Rating of Health Information on the Net with MedCERTAIN: Using a medPICS/RDF/XML metadata structure for implementing eHealth ethics and creating trust globally. Journal of Medical Internet Research. 2000; 2 (3/supp.2): e1. Online (www.jmir.org/2000/3/suppl2/e1/index.htm). Accessed November 18, 2000.

22. Foltz A, Sullivan J. Reading level, learning presentation preference, and desire for information among cancer patients. J Cancer Educ. 1996 Spring; 11(1):32-8.

23. Fox S, Rainie L. The online health care revolution: How the Web helps Americans take better care of themselves. Washington, DC: Pew Charitable Trusts, 2000.

24. Fox S, Horrigan J, Lenhart A, et al. More online, doing more: 16 million newcomers gain Internet access in the last half of 2000 as women, minorities, and families with modest incomes continue to surge online. Washington, DC: Pew Internet & American Life Project, February 2001.

25. Fry E. The Readability graph validated at primary levels. The Reading Teacher. 1969;22: 534-538.

26. Fry E. Fry's readability graph: Clarifications, validity, and extension to level 17. J Reading. 1977;21(3):242-252.

27. Gilliam B, Pena S, Moutain L. The Fry graph applied to Spanish readability. The Reading Teacher. January, 1980:426-430.

28. Ginzberg E. Access to health care for Hispanics. JAMA. 1991;265(2):238-41.

29. Gorman C. Doctors Without Borders. On Magazine. April 2001: 36-37.

30. Griffiths KM, Christensen H. Quality of web based information on treatment of depression: Cross sectional survey. BMJ 2000;321:1511-1515.

31. Health on the Net Foundation. HON code on conduct (HONcode) for medical and health web sites. Health on the Net Foundation. Online (http://www.hon.ch/ HONcode). Accessed June 13, 2000.

32. Hegstrom E. Gore and Bush employ splintered Spanish: But Hispanics seem to applaud the effort. Houston Chronicle. September 27, 1999. Online (http://seattlepi.nwsource. com/national/span27.shtml). Accessed January 10, 2000.

33. Hibbard J, Slovic P, Jewett J. Informing consumer decisions in health care: Implications from decision-making research. The Milbank Quarterly. 1997;75(3): 395-414.

34. Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the world wide web: systematic survey of advice on managing fever in children at home. BMJ 1997;314:1875-1879.

35. Internet.com Corp. Search Engine Watch. How search engines rank Web pages. Online (http://www.searchenginewatch.com). Accessed January 5, 2001.

36. Internet Healthcare Coalition. Tips for Health Consumers: Finding Quality Health Information on the Internet. Pennsylvania: Internet Healthcare Coalition, 2000. Online (http:// www.ihealthcoalition.org). Accessed December 18, 2000.

37. Jiang YL. Quality evaluation of orthodontic information on the World Wide Web. Am J Orthod Dentofacial Orthop 2000;118:4-9.

38. Johnston P. Reading Comprehension Assessment: A Cognitive Basis. Newark, DE: International Reading Association; 1983.

39. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Arch Gen Psychiatry. 51:8–19. 1994.

40. Kiefer, KM. Health Literacy: Responding to the Need for Help. Washington, DC: Center for Medicare Education; 2001.

41. Kirsch IS, Jungeblut A, Jenkins L, Kolstad A. Adult literacy in America. Washington, DC: National Center for Education Statistics; 1993.

42. Klare G. Assessing readability. Reading Research Quarterly. 1974;1:62-102.

43. Landro L. Web Health Groups Ponder How to Set Universal Standards. Wall Street Journal. November 3, 2000; p1.

44. Linde K, Mulrow CD. St. John's Wort for depression [Cochrane Review on CD-ROM]. Oxford, England; Cochrane Library, Update Software; 1998:issue 4.

45. Lopez DE. Language and assimilation. In: Waldinger R, Bozorgmehr M, eds., Ethnic Los Angeles. New York, NY: Russell Sage Foundation; 1996:139-163.

46. Mannino, Homa, Pertowski, et al. Surveillance for Asthma: United States, 1960-1995. Morbid Mortal Wkly Rep. 1998;47:1-28.

47. Mayberry RM, Mili F, Ofili E. Racial and Ethnic Differences in Access to Medical Care. Med Care Res Rev. 2000; 57(suppl 1):108-145.

48. McClung HJ, Murray HD, Heitlinger LA. The internet as a source for current patient information. Pediatrics 1998;101:1-4.

49. Metametrics, Inc. The Lexile Framework for Reading. Online (http://www.lexile.com). Accessed December 11, 2000.

50. Mokdad AH, Serdula MK, Dietz WH, Bowman BA, Marks JS, Koplan JP. The spread of the obesity epidemic in the United States. JAMA. 1999;282:1519-1522

51. Mulrow CD, Williams JW, Trivedi M, et al. *Treatment of Depression: Newer Pharmacotherapies*. Rockville, MD: Agency for Health Care Policy and Research; 1999.

52. Mynors-Wallis LM, Gath DH, Lloyd-Thomas AR, Tomlinson D. Randomised controlled trial comparing problem solving treatment with amitriptyline and placebo for major depression in primary care. BMJ. 1995; 310:441-445.

53. National Heart, Lung and Blood Institute. Clinical guidelines on the identification, evaluation and treatment of overweight and obesity in adults. Bethesda, MD: National Institutes of Health; 1998. NIH Publication #98-4083.

54. National Task Force on the Prevention and Treatment of Obesity. Dieting and the development of eating disorders in overweight and obese adults. Arch Intern Med 2000; 160: 2581-2589.

55. Notess GR. Search Engines Statistics: Database Overlap. Search Engine Showdown, 2000. Online (http://searchengineshowdown.com/stats/overlap.shtml). Accessed December 7, 2000.

56. Ostrom CM. Internet medicine opens can of worms: tremendous potential for good – and harm. *Seattle Times*. May 12, 1999;A1.

57. Overland JE, Hoskins PL, McGill MJ, Yue DK. Low literacy: a problem in diabetes education. Diabet Med. 1993;10(9):847-50.

58. Pear R. U.S. and Thai Officials Attack Sales of Medicines on Internet. *New York Times*. March 21, 2000, A1.

59. Rippen H, Risk A. E-Health code of ethics. Journal of Medical Internet Research. 2000; 2 (2): e9. Online (http://www.jmir.org/2000/2/e9). Accessed June 13, 2000.

60. Romano R. Readability in Cancer Communications: Methods, Examples and Resources for Improving the Readability of Cancer Messages and Materials. Bethesda, MD: U.S. Department of Health, Education and Welfare, Public Health Service, National Institutes of Health, National Cancer Institute; 1979.

61. Root J, Stableford S. Easy to read consumer communication: A missing link in medicaid managed care. J Health Politics Policy Law. 1999;24(1): 1-26.

62. Sandvik, H. Health Information and Interaction on the Internet: A Survey of Female Urinary Incontinence. BMJ. 1999;319:32-6.

63. Search Engine Guide. Search Engine Guide. Online (http://www.searchengineguide.com). Accessed January 5, 2001.

64. Smith DH, Malone DC, Lawson KA, Okamoto LJ, Battista C, Saunders WB. A National Estimate of the Economic Costs of Asthma. Am J Respt Crit Care Med. 1997;156:787-793.

65. Smith JP, Edmonston B. The face of the U.S. population in 2050. In: Smith JP, Edmonston B, eds. The New Americans (pp. 76-134). Washington, DC: National Academy Press, 1997.

66. Soot LC, Moneta GL, Edwards JM. Vascular surgery and the internet: a poor source of patient-oriented information. J Vasc Surg 1999;30:84-91.

67. Sullivan D. How Search Engines Work. Search Engine Watch, 2000. Online (http:// searchenginewatch.com/webmasters/work.html). Accessed July 4, 2000. Sullivan, D.

68. Taylor H. Cyberchondriacs Update. Harris Poll #19. New York, NY: Harris Interactive; April 2001.

69. Taylor H. Explosive Growth of "Cyberchondriacs" Continues. Online (http://www. louisharris.com/harrispoll/index.asp?PID=104). Accessed January 5, 2001.

70. United States Bureau of Commerce. Falling Through the Net: Defining the Digital Divide. 1999. Online (http://www.ntia.doc.gov/ntiahome/fttn99/contents.html). Accessed January 11, 2001.

71. United States Census Bureau. Detailed Language Spoken at Home and Ability to Speak English for Persons 5 Years and Over – 50 Languages with Greatest Number of Speakers: U.S. 1990. Online (http://www.census.gov/population/socdemo/ language/table5.txt). Accessed December 19, 2000.

72. United States Census Bureau. Population by race and Hispanic or Latino origin for the United States: 1990 and 2000. April 2, 2001. Online (http://www.census.gov/ population/cen2000/phc-t1/tab04.pdf). Accessed: April 30, 2001.

73. United States Department of Health and Human Services. Writing and Designing Print Materials for Beneficiaries: A Guide for States Medicaid Agencies. HCFA Publication Number 10145; 1999.

74. Williams DM, Counselman FL, Caggiano CD. Emergency department discharge instructions and patient literacy: a problem of disparity. Am J Emerg Med. 1996 Jan; 14(1):19-22.

75. U.S. Preventive Services Task Force. Screening for breast cancer. Guide to clinical preventive services. Washington, DC: Dept of Health and Human Services; 1996.

76. Wilson K, Mottram P, Sivanthan A. A review of antidepressant drug trials in the treatment of older depressed people [protocol for a Cochrane Review]. Oxford, England: Cochrane Library, Update Software; 1998:issue 4.

77. Wilson FL. Measuring patients' ability to read and comprehend: a first step in patient education. Nursing Connections. 1995; 8(4):17-25.

78. Wright BD, Strenner AJ. Readability and Reading Ability. Presentation to Australian Council on Education Research (ACER); 1998.