JAMA

CURRENT ISSUE   INDEXES   PAST ISSUES

Original Contribution

Health Information on the Internet

Accessibility, Quality, and Readability in English and Spanish

Gretchen K. Berland, MD; Marc N. Elliott, PhD; Leo S. Morales, MD, PhD; Jeffrey I. Algazy, MD, MPH; Richard L. Kravitz, MD, MSPH; Michael S. Broder, MD; David E. Kanouse, PhD; Jorge A. Muñoz, MS, MPhil; Juan-Antonio Puyol; Marielena Lara, MD, MPH; Katherine E. Watkins, MD, MSHS; Hannah Yang; Elizabeth A. McGlynn, PhD

Context Despite the substantial amount of health-related information available on the Internet, little is known about the accessibility, quality, and reading grade level of that health information.

Objective To evaluate health information on breast cancer, depression, obesity, and childhood asthma available through English- and Spanish-language search engines and Web sites.

Design and Setting Three unique studies were performed from July 2000 through December 2000. Accessibility of 14 search engines was assessed using a structured search experiment. Quality of 25 health Web sites and content provided by 1 search engine was evaluated by 34 physicians using structured implicit review (interrater reliability >0.90). The reading grade level of text selected for structured implicit review was established using the Fry Readability Graph method.

Main Outcome Measures For the accessibility study, proportion of links leading to relevant content; for quality, coverage and accuracy of key clinical elements; and grade level reading formulas.

Results Less than one quarter of the search engine's first pages of links led to relevant content (20% of English and 12% of Spanish). On average, 45% of the clinical elements on English- and 22% on Spanish-language Web sites were more than minimally covered and completely accurate and 24% of the clinical elements on English- and 53% on Spanish-language Web sites were not covered at all. All English and 86% of Spanish Web sites required high school level or greater reading ability.

Conclusion Accessing health information using search engines and simple search terms is not efficient. Coverage of key information on English- and Spanish-language Web sites is poor and inconsistent, although the accuracy of the information provided is generally good. High reading levels are required to comprehend Web-based health information.

The Internet is an increasingly important source of health-related information for consumers. One recent survey estimated that more than 60 million US residents went online in search of health information in the past year.[1] The online population is becoming more representative of the larger US population in terms of race, age, income, and educational attainment.[2] Among those who use the Internet, more than 70% report the health information they find influences a decision about treatment.[1]

The ability to obtain accurate medical information quickly, conveniently, and privately online presents to consumers an opportunity for better-informed decision making and greater participation in care.[3] Little is known, however, about whether the available material is sufficiently complete and accurate to support consumer decision making. Several studies of single medical conditions have suggested deficiencies in the quality of Web-based health information.[4-11]

Several organizations have developed criteria to guide and evaluate health-related Web site content (eg, HON Code, American Medical Association, Internet HealthCare Coalition, Hi-Ethics, MedCertain),[12-18] but these criteria have not been systematically applied to a broad set of Web pages and conditions. Furthermore, because many of these systems rely on voluntary self-assessments by Web page developers, the reliability and validity of many of these evaluations is unknown.[19, 20]

Even if online materials are comprehensive and accurate, the ability of users to apply these assessment tools depends on their ability to locate and understand those materials. The Internet has the potential to eliminate barriers in access to information for patients, but only if online material can be read and understood by many different types of users.[21, 22]

Preliminary data from the 2000 US Census indicate that the population is becoming increasingly diverse. Since 1990, the US Hispanic population has grown from 22.3 million to 35.3 million, making Hispanics the largest minority group in the United States.[23] Among immigrant Hispanics, more than 98% report speaking primarily Spanish at home.[24] 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 face greater barriers to traditional sources of medical care and information.[25, 26] We are unaware of any studies that have evaluated Spanish-language materials.

We conducted a large cross-sectional study to describe and evaluate health information on the Internet in English and Spanish. We evaluated information that we found using search engines and by visiting health-related Web sites on 4 medical conditions: breast cancer, childhood asthma, depression, and obesity. We asked 4 main questions: What are consumers likely to find when they search online about these conditions? How comprehensive is the information? How accurate is it? At what grade reading level is the material presented?

METHODS

We conducted 3 studies to assess the accessibility of relevant content; the quality of health information; and the reading grade level of text. Each study used different methods to assess the same 4 conditions. Conditions were selected by project staff based on prevalence, clinical significance, and diversity of the affected populations.[27-41] Each study was conducted independently in English and Spanish.

Study 1: Accessibility of Relevant Content

Selecting Search Engines
Search engines are designed to help people locate information on the Internet. To assess how well search engines perform this function, we selected 10 English-language and 4 Spanish-language search engines. Three of the English-language and 2 of the Spanish-language search engines were chosen based on popularity (defined as the number of unique visitors per month as reported by Media Metrix, Inc in June 2000[42]). The remaining 9 search engines were selected because they featured unique methods of ranking Web sites.[43] Examples of ranking methods included ranking by location and frequency of key words within a site; ranking by the number of times a site is linked to by another site; ranking by payment from sites; and ranking by human editing.[43], [44]

Conducting Standardized Searches
Trained searchers entered the 4 search terms ("breast cancer," "childhood asthma," "depression," and "obesity") into each of the 14 search engines. All links on the first electronic page for each search engine were then counted and classified. Links were classified as relevant if the search term or 1 of 30 to 40 related key terms per condition (eg, tamoxifen, inhaler, gastric bypass surgery, St John's wort) was present in the link itself or the surrounding text.

Searchers then followed a sample of relevant links to determine whether they led to relevant content. One sample was the first 5 relevant links on the search results page. All remaining links were enumerated and divided into 5 strata of equal size; 1 relevant link from each stratum was selected randomly. Searchers clicked on 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 site). If the first relevant link led to a content page, the page was saved for further analysis. If the first link led to more links, the searcher randomly selected a relevant link from the first 15 relevant links on that Web page. If searchers had not reached a content page after 10 cycles, the search was discontinued.

Characterizing Content
Using a standardized form, trained coders first classified Web page content by relevance. Web pages were coded as relevant if they contained any materials related to the 4 search terms or the 30 to 40 key terms related to each condition. Coders then assessed the relevant pages for the presence of promotional content (defined as material designed to encourage site visitors to purchase products or services or participate in research programs sponsored by the site). Explicit advertisements were classified separately from promotional material and had to be located in a banner or

sidebar on the Web page.

Study 2: Quality of Health Information on the Internet

Selecting Web Sites
Eighteen unique English-language health Web sites (6 general health, 12 condition-specific) and 7 unique Spanish-language Web sites (3 general health, 4 condition-specific) were selected for this study (Table 1). We selected 6 English-language general health Web sites that were ranked highly in 2 widely used Internet industry reports, Cyber Dialogue and PC Data Online for September 2000.[45], [46] Content provided by one of the most popular search engines was also included.[42] 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. Project staff limited Web sites to those not requiring subscriptions or payments.

Developing Condition-Related Topics and Questions
Panels of 3 to 4 nationally recognized clinical experts and representatives from patient advocacy organizations identified 5 to 8 key clinical topic areas for each condition (26 in all). Panelists were recruited for their clinical or scientific experience, familiarity with national guidelines, current research, or national reputations in the medical conditions of interest. No panelist had consulted for, or had any financial involvement with, any e-health Web site. Panelists were asked to identify topics that were relevant to patients, their families, or laypersons seeking information on the study conditions. Panelists also considered whether it was reasonable to expect to find this information on the Web. The panels then wrote 36 consumer-oriented questions relating to the 26 topics.

For example, the topic "breast cancer screening" was characterized by the following questions: "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?" A complete list of all condition-related topics and questions is located in Online Table 1 available in PDF format joc02274t1.pdf.

Development of Clinical Elements
To enhance the consistency of the structured implicit review, the 4 clinical panels each developed a series of 1 to 8 clinical elements for each of the questions based on evidence-based guidelines and materials from selected literature reviews.[27-41] For example, for the topic of breast cancer screening, 4 clinical elements were developed. These included the following: women older than 50 years 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 of the disease; and a lack of consensus exists about the need for or appropriate interval of mammography in women from age 40 to 49 years. A total of 100 clinical elements were developed (Online Table 1 available in PDF format joc02274t1.pdf).

Retrieving Health Information
Four abstractors (2 monolingual in English, 2 bilingual in English and Spanish) independently reviewed each Web site (spending a maximum of 90 minutes per site using high-speed Internet connections) on October 18-30, 2000, and November 6-13, 2000, to retrieve content related to the

questions (Online Table 1 available in PDF format joc02274t1.pdf).
Abstractors did not receive any of the condition-related clinical elements prior to conducting each search. On average, 65% of retrieved Web pages were common between abstractors. Search results were saved using a software application (CatchTheWeb, Math Strategies, Greensboro, NC) that enabled project researchers to accurately save, abstract, and manage Web pages for later use.

Retrieved materials were stripped of identifying information, printed, and assembled into notebooks. Each notebook contained the materials retrieved from a single search on a Web site (eg, 1 condition per site). The 78 unique English-language notebooks averaged 250 printed pages (range, 21-547 printed pages). The 32 unique Spanish-language notebooks averaged 68 printed pages (range, 8-366 printed pages). A total of 21 711 printed pages (2660 Web pages, defined by the programmer's end-of-page mark) were abstracted across 4 conditions: 19 529 printed pages (2262 Web pages) from English-language and 2182 printed pages (398 Web pages) from Spanish-language Web sites.

Evaluating the Web Sites
Thirty-four physicians (30 monolingual in English, 4 bilingual in English and Spanish) from around the United States were recruited to evaluate the abstractor-retrieved material. All 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. No reviewer rated more than 5 notebooks for any condition or evaluated materials from the same Web site twice. Forty English-language (51%) and 14 Spanish-language (44%) randomly selected notebooks were evaluated by 2 reviewers. Each Web site underwent 2 to 4 reviews per condition.

Four standardized rating forms were developed that listed the condition-related topics, questions, and clinical elements (eg, 1 condition per form). Reviewers were asked to rate the level of coverage for each clinical element as not addressed, minimally addressed, or more than minimally addressed. *Not addressed* meant there was no reference to the issue on any page of the notebook. *Minimally addressed* meant the clinical element was mentioned at least briefly. For example, for breast cancer screening, if mammography was mentioned as a way to identify early breast cancer, 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 was considered minimal coverage. *More than minimally addressed* meant that most of the clinical elements were mentioned and the level of explanation was more than cursory. For example, coverage was considered 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 years, or that breast cancer may be detected earlier by mammography than physical examination, or if a detailed discussion of the pros and cons of mammography and the appropriate ages for screening were provided.

Reviewers also rated the accuracy of content for each clinical element that was at least minimally addressed: mostly incorrect, mostly correct, and completely correct.

After rating Web site materials on coverage and accuracy, reviewers were asked to list instances of conflicting information found during their review. These conflicts were not limited to the set of clinical elements for which

coverage and accuracy were evaluated. Six categories of conflicting information were identified: (1) treatments; (2) diagnosis; (3) definitions; (4) adverse effects; (5) etiology and risk factors; and (6) incidence and prevalence. Two project physicians (R.L.K. and J.I.A.) 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 physicians were included in the final analysis.

Analytic Methods

We used Stata statistical software (version 6.0; Stata Corporation, College Station, Tex). The unit of analysis was the link (specific URL [uniform resource locator]) for the study of search engine efficiency, the standardized rating form for the study of quality, and the Web site for the study of grade reading level.

Rating forms contained multiple ratings (corresponding to clinical elements) of coverage and accuracy using the 3-point ordinal scales mentioned previously. For purposes of analysis, summary measures were computed by averaging across elements within a given rating form.

All analyses were conducted separately for English- and Spanish-language search engines and Web sites. All statistical tests were 2-sided and were assessed for significance at the .05 level. Measures were tested for variation by condition, search engine, and site, as applicable. A 2-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 established that variation in the outcome of interest was statistically significant for the categorical variable (condition, search engine, site), a series of 2-sample follow-up tests were performed comparing the outcome at each level of the categorical variable with the overall distribution of the outcome.

The omnibus tests used were 1-way analysis of variance, the Kruskal-Wallis rank-sum test, and the x^2 test of homogeneity for measures that were normally, ordinally, and nominally distributed, respectively. Two-sample *t* tests, Wilcoxon rank-sum tests, and x^2 tests of homogeneity were the corresponding follow-up tests.

Interrater Reliability

In the search engine study, interrater reliability of the judgments by searchers and coders was high for both classification of links and content ($\kappa \geq 0.80$).[47]

In the Web site study, 2 measures of interrater reliability of Web site reviewers were computed. A standard measure of reliability, computed as the correlation in ratings between reviewers examining identical notebooks of material retrieved from the same Web site, was calculated. To assess the sensitivity of reviewer ratings to variation in the retrieved material (eg, the material retrieved by abstractor 1 vs abstractor 2 on the same Web site for the same condition), a second, more stringent measure of reliability was computed as the correlation in ratings between reviewers examining different notebooks of material from the same Web site and condition. We computed 16 interrater reliabilities by the standard rule and 16 by the stringent rule for each language: 1 for every combination of the 4 conditions and the 4 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 interrater reliability on English-language Web sites and 12 reviews were included in each calculation of interrater reliability on Spanish-language Web sites. The standard interrater reliability was 0.90 or greater for all conditions and measures, averaging 0.96 for both English- and Spanish-language sites. The second measure of interrater reliability averaged 0.77 for English- and 0.60 for Spanish-language Web sites.

Study 3: Reading Grade Levels of Web Sites

To determine reading grade level, we used the Fry Readability Graph (FRG) method, which has been validated in both English and Spanish.[48], [49] Three sample passages of text exactly 100 words in length from the beginning, middle, and end of the material abstracted from each Web site were selected. For each 100-word sample, the number of sentences and syllables were counted. The FRG calculates an estimated grade level as a function of the average number of sentences and average number of syllables for each source document. The FRG accounts for the fact that Spanish documents tend to have more syllables per word than English documents of the same reading level.[50]

RESULTS

Efficiency of Searches and Type of Information

Figure 1 and Table 2 summarize the experience that someone seeking information would have when using a search engine. The first page of search results from all English-language search engines listed 3735 links, 1265 (34%) of which were relevant. The proportion of these links that were relevant varied significantly by search engine (*P*<.001). Among 389 sampled relevant links, 288 (74%) selected led to a content page within 10 clicks, and 230 (79%) of those pages contained content relevant to the search topic. Thus, when following apparently relevant links, relevant content was identified 59% of the time (Figure 1 and Table 2). There was significant variation in the likelihood of reaching relevant content from potentially relevant links by search engine (range, 35%-88%, *P*<.001, Table 2). One in 5 (20%) links on the first page of search results led to relevant content (Table 2). There was no significant variation among search engines in the probability that first-page links would lead to relevant content.

Results for Spanish-language search engines were similar. The first page of results returned 1685 links, 296 (18%) of which were relevant. Among the 151 selected relevant links, 101 (67%) led to content and 95 (94%) of those pages contained relevant content. Overall, 63% of relevant links led to relevant content (Figure 1 and Table 2). There was significant variation in the likelihood of reaching relevant content by search engine (range, 49%-78%, *P*<.001, Table 2). Twelve percent of all links on the first page of search results led to relevant content, with no significant variation by search engine (Table 2).

Fifty-six percent (n = 129) of the relevant English-language content pages contained explicit advertisements and 44% (n = 101) contained other promotional material. The presence of advertisements and promotional materials on relevant Spanish-language content pages was 36% (n = 34) and 21% (n = 20), respectively.

Quality of Health Information

Coverage of Topics

Coverage is reported as the mean proportion of clinical elements across sites with no coverage; minimal coverage; and more than minimal coverage of the clinical elements for each condition. Among English-language sites, the mean percentage of clinical elements that were not covered varied significantly across conditions: 16% for breast cancer, 27% for childhood asthma, 20% for depression, and 35% for obesity (Table 3). Topics that were not covered most often included alternatives to standard medical and surgical treatments for breast cancer (28%), symptoms suggestive of poorly controlled asthma (48%), evaluation of depression (33%), and safety and effectiveness of dietary supplements used for obesity (61%).

On Spanish-language Web sites, the mean percentage of clinical elements receiving no coverage also varied significantly across conditions: 49% for breast cancer, 33% for childhood asthma, 61% for depression, and 69% for obesity (Table 3). Topics that were not covered most often included alternatives to standard medical and surgical treatments for breast cancer (90%), expected benefits and possible adverse effects of asthma therapies (44%), evaluation of depression (84%), safety and effectiveness of dietary supplements (100%), and types of popular diets for obesity (100%).

Accuracy of Information

On English-language Web sites, the mean percentage of covered clinical elements for which the text was completely correct was 91% for breast cancer, 84% for childhood asthma, 75% for depression, and 86% for obesity. In Spanish, the mean percentages were 96% for breast cancer, 53% for childhood asthma, 63% for depression, and 68% for obesity.

On English-language sites, the mean percentages of covered clinical elements rated as mostly incorrect were 0% for breast cancer, 3% for childhood asthma, 3% for depression, and 3% for obesity. In Spanish, the mean proportions were 0% for breast cancer, 4% for childhood asthma, 18% for depression, and 0% for obesity. As an example, one depression site stated that omega-3 fatty acid deficiencies cause major depressive disorders. One childhood asthma site describes cockroaches as the leading cause of asthma among children.

Combined Measure of Coverage and Accuracy

In English, the mean percentage of clinical elements receiving more than minimal coverage that were completely accurate was 63% for breast cancer, 36% for childhood asthma, 44% for depression, and 37% for obesity. For breast cancer, depression, and obesity, there was significant variation among English-language Web sites (Table 3). Two sites performed statistically better than average: http://www.Oncolink.com for breast cancer and http://www.nimh.nih.gov for depression (for both, *P* = .02). No Web site was statistically better than the condition average for childhood asthma and obesity.

On Spanish-language Web sites, the corresponding proportions receiving

more than minimal coverage that were completely accurate were 39% for breast cancer, 23% for childhood asthma, 12% for depression, and 15% for obesity. There was significant variation among Web sites for breast cancer and depression (*P*<.05), but no Web site was statistically better than the condition average.

For a comprehensive summary of coverage and accuracy of elements of condition-related topics for the 4 conditions, see Online Table 2 available in PDF format joc02274t2.pdf.

Conflicting Information
Overall, just over half of English-language Web site reviews revealed 1 or more conflicts in the information provided (Table 4; Spanish reviewers noted no conflicts). Conflicts involved treatment (present in 35% of reviews), diagnosis (13%), definitions (7%), adverse effects (5%), etiology and risk factors (5%), and incidence and prevalence (4%). As an example, a childhood asthma site stated at one point that inhaled steroids do not stunt growth and later stated that inhaled steroids do stunt growth. Materials on depression were the most likely to have conflicts on treatment, whereas breast cancer materials were the most likely to contain conflicts on diagnosis (Table 4, *P*<.001).

Reading Grade Level

For English-language Web sites, the average reading level was collegiate (mean [SD] grade, 13.2 [2.1]) and ranged from 10th grade to graduate school level (Figure 2). For the Spanish-language Web sites, the average reading level was at 10th grade (mean [SD] grade, 9.9 [2.5]) and ranged from grades 7 to 13 (Figure 2). The mean grade reading level for the English-language Web sites was significantly higher than for Spanish-language Web sites (*P*<.003).

COMMENT

To our knowledge, this is the first study to examine English- and Spanish-language health information on the Internet across multiple conditions. We found that search engines are only moderately efficient in locating information on a particular health topic. 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 (Carolyn Gratzer, Cyber Dialogue, oral communication, October 13, 2000). Overall, 1 in 5 links identified by 10 English-language and 1 in 8 links from 4 Spanish-language search engines led to a Web page with relevant content.

We examined 2 dimensions of Web site quality: whether key consumer questions were covered and whether the information was accurate. Although we found thousands of pages of material related to key questions, there were substantial gaps in the availability of key information. Only half of the topics that the expert panels thought were important for consumers were covered more than minimally. This deficiency was particularly striking

across Spanish-language sites, where more than half of the condition-related topics were not addressed.

Our results suggest that consumers using the Internet may have a difficult time finding complete and accurate information on a health problem. If people are relying on the Internet to make treatment decisions, including whether to seek care, deficiencies in information could negatively influence consumer decisions. 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.

The reading level of most Web-based material is quite high. All of the English-language sites had material that required at least a 10th-grade reading level, and more than half of the sites presented material at the college level. Although 1 Spanish-language site presented material at the elementary school level, all others required at least a ninth-grade reading level. According to the 1992 National Adult Literacy Survey, 92 million adults in the United States — almost 48% of the population — and more than 75% of current welfare recipients have low or very low reading skills.[51] Thus, even if wider access to computer technologies narrowed the digital divide, the online health information currently available would be difficult for many people to understand.

This study has some important limitations. First, the Internet is a moving target, and we were able to take only a snapshot of its performance. Changes in content over time are not represented. However, without dedicated attention, it seems unlikely that the variability in performance, gaps in availability of information, and high reading levels will change dramatically. Second, we looked at a small set of search engines, Web sites, and conditions, and hence cannot draw more general conclusions. However, because we included the most popular search engines and Web sites, the results are likely to reflect common experiences. Their variability in performance suggests that the likelihood of finding the information one needs, on the topic of one's choice, will depend on where one starts. Third, we studied the performance of search engines using very simple search terms. Had more sophisticated search strategies been used, our findings might have been different. Fourth, our research was not a natural experiment (eg, using actual consumers to search for information and testing their knowledge after such a search), so we cannot draw conclusions about what people actually encounter when they search for information, or about how well they are able to interpret the information they find. Fifth, the necessary inclusion of medical terms in analyzed text may be partially responsible for high estimated grade reading levels, although we assessed Flesch-Kincaid scores[52] in the same passages with and without the medical terminology included, and when medical terminology was removed, the grade level declined by only 0.3 grade levels on average (range, 0.1-0.8). Sixth, because the Internet and many Web sites make available a large volume of material, it is possible that our searchers missed information that was available on a site. For that reason, we had 2 searchers look for information on each site, and they found different material. But the conclusions that reviewers reached about sites were quite consistent, even when the retrieved material they evaluated was different. Furthermore, our searchers were skilled, trained for the task, and devoted more time to finding information than people report spending on average. Thus, if our searchers could not find the information, probably most consumers also would have difficulty doing so.

Our results suggest several ways to make Web-based information more useful. First, variation among search engines suggests that it is possible to improve search efficiency, perhaps by improving the methods for indexing Web pages. Second, the lack of critical information for each of the 4 conditions suggests that Web site developers should focus on providing more complete information. Third, Web site developers need to ensure that the information is accurate and free from conflict. Although accuracy levels were generally high, the presence of conflicting information makes it possible that people will be more confused than enlightened. Fourth, some mechanism for routinely rating Web sites for coverage and accuracy may be useful. Comprehensive assessments of the type conducted for this study are highly labor intensive, but simpler methods also may be effective. Fifth, information on the Web needs to be made more readable if the Internet is to serve as a "leveler" across different socioeconomic backgrounds.

The Internet has the potential to be a powerful resource for meeting some of the public's health information needs. Ideally, consumers would be able to learn much of what they need to know from high-quality Web sites, so that the limited time they have with their physicians could be used more efficiently. However, this requires that Web sites present well-organized and accurate information in a way that is understandable. Research is needed on how the public's use of the Internet facilitates, complements, or complicates patient-physician communication and on how patients and health professionals can make better use of this resource.

Author/Article Information

Author Affiliations: RAND Health (Drs Berland, Elliott, Morales, Algazy, Broder, Kanouse, Lara, Watkins, and McGlynn, Mr Puyol, and Ms Yang) and RAND Graduate School (Mr Muñoz), Santa Monica, Calif; Robert Wood Johnson Clinical Scholars Program, University of California, Los Angeles (Drs Berland and Algazy); University of California, Los Angeles School of Medicine (Drs Morales, Broder, and Lara); and University of California, Davis School of Medicine (Dr Kravitz).

Corresponding Author and Reprints: Gretchen K. Berland, MD, 1700 Main St, M-29, Santa Monica, CA 90407-2138 (e-mail: berland@rand.org).

Author Contributions: *Study concept and design:* Berland, Elliott, Morales, Kravitz, Algazy, Broder, Kanouse, Puyol, Lara, Watkins, Yang, McGlynn.

Acquisition of data: Berland, Morales, Algazy, Muñoz, Puyol, Yang.

Analysis and interpretation of data: Berland, Elliott, Morales, Kravitz, Algazy, Kanouse, Muñoz, McGlynn.

Drafting of the manuscript: Berland, Elliott, Morales, Kravitz, Algazy, Broder, Kanouse, Puyol, Watkins, Yang, McGlynn.

Critical revision of the manuscript for important intellectual content: Berland, Elliott, Morales, Kravitz, Algazy, Broder, Kanouse, Muñoz, Lara, McGlynn.

Statistical expertise: Elliott, Morales, Muñoz.

REFERENCES

1.
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.

2.
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.

3.
Goldsmith J.
How will the Internet change our health system?
Health Aff (Millwood).
2000;19:148-156.
MEDLINE

4.
Beredjiklian PK, Bozentka DJ, Steinberg DR, Bernstein J.
Evaluating the source and content of orthopedic information on the Internet:
the case of carpal tunnel syndrome.
J Bone Joint Surg Am.

2000 Nov;82-A:1540-1543.

5.
Biermann JS, Golladay GJ, Greenfield MLVH, Baker LH.
Evaluation of cancer information on the Internet.
Cancer.
1999;86:381-390.
MEDLINE

6.
Davison K.
The quality of dietary information on the World Wide Web.
Clin Perform Qual Health Care.
1997;5:64-66.
MEDLINE

7.
Griffiths KM, Christensen H.
Quality of web based information on treatment of depression: cross sectional survey.
BMJ.
2000;321:1511-1515.
MEDLINE

8.
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.
MEDLINE

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

10.
McClung HJ, Murray HD, Heitlinger LA.
The Internet as a source for current patient information.
Pediatrics.
1998;101:1-4.
MEDLINE

11.
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.
MEDLINE

12.
Health on the Net Foundation.

HON code on conduct (HONcode) for medical and health web sites. Available at: http://www.hon.ch/HONcode/. Accessed June 13, 2000.

13.
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.
J Med Internet Res.
2000;2(3/suppl 2):e1. Also available at: http://www.jmir.org/2000/3/suppl2/index.htm. Accessed November 18, 2000.

14.
Ambre J, Guard R, Perveiler FM, Renner J, Rippen H.
Criteria for assessing the quality of health information on the Internet.
May 4, 1999. Available at: http://hitiweb.mitretek.org/docs/criteria.html. Accessed August 11, 2000.

15.
Rippen H, Risk A.
E-health code of ethics.
J Med Internet Res.
2000;2:e9. Also available at: http://www.jmir.org/2000/2/e9. Accessed June 13, 2000.

16.
Tips for Health Consumers: Finding Quality Health Information on the Internet.
Washington, DC: Internet Healthcare Coalition; 2000. Also available at: http://www.ihealthcoalition.org/. Accessed December 18, 2000.

17.
Winker MA, Flanagin A, Chi-Lum B, et al.
Guidelines for medical and health information sites on the Internet.
JAMA.
2000;283:1600-1606. Also available at: http://jama.ama-assn.org/issues/v283n12/ffull/jsc00054.html. Accessed April 23, 2001.
ABSTRACT | FULL TEXT | PDF | MEDLINE

18.
Health Internet ethics: ethical principles for offering Internet health services to consumers.
May 2000. Available at: http://www.hiethics.com/principles/index.asp. Accessed April 23, 2001.

19.
Kim P, Eng TR, Deering MJ, Maxfield A.
Published criteria for evaluating health related Web sites: review.

BMJ.
1999;318:647-649.
MEDLINE

20.
Baur C, Deering MJ.
Proposed frameworks to improve the quality of health Web sites: review.
Med Gen Med.
2000;:E35.

21.
Edejer TT.
Disseminating health information in developing countries: the role of the Internet.
BMJ.
2000;321:797-800.
MEDLINE

22.
McLellan F.
"Like hunger, like thirst": patients, journals, and the Internet.
Lancet.
1998;:SII39-SII43.

23.
US Census Bureau.
Population by race and Hispanic or Latino origin for the United States: 1990 and 2000.
April 2, 2001. Available at:
http://www.census.gov/population/cen2000/phc-t1/tab04.pdf. Accessed April 30, 2001.

24.
Lopez DE.
Language and assimilation.
In: Waldinger R, Bozorgmehr M, eds. *Ethnic Los Angeles*. New York, NY: Russell Sage Foundation; 1996:139-163.

25.
Ginzberg E.
Access to health care for Hispanics.
JAMA.
1991;265:238-241.
MEDLINE

26.
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.
MEDLINE

27.

Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB.
Annual deaths attributable to obesity in the United States.
JAMA.
1999;282:1530-1538.
ABSTRACT | FULL TEXT | PDF | MEDLINE

28.
Mokdad AH, Serdula MK, Dietz WH, Bowman BA, Marks JS, Koplan JP.
The spread of the obesity epidemic in the United States, 1991-1998.
JAMA.
1999;282:1519-1522.
ABSTRACT | FULL TEXT | PDF | MEDLINE

29.
Mannino DM, Homa DM, Pertowski CA, et al.
Surveillance for asthma: United States, 1960-1995.
Mor Mortal Wkly Rep CDC Surveill Summ.
1998;47:1-27.
MEDLINE

30.
Smith DH, Malone DC, Lawson KA, et al.
A national estimate of the economic costs of asthma.
Am J Respir Crit Care Med.
1997;156:787-793.
MEDLINE

31.
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.
1994;51:8-19.
MEDLINE

32.
Winchester DP, Cox JD.
Standards for diagnosis and management of invasive breast carcinoma.
CA Cancer J Clin.
1998;48:83-107.
MEDLINE

33.
US Preventive Services Task Force.
Screening for Breast Cancer: Guide to Clinical Preventive Services.
Washington, DC: US Dept of Health and Human Services; 1996.

34.
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.
ABSTRACT | FULL TEXT | PDF | MEDLINE

35.

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.

[36.](#)
Mulrow CD, Williams JW, Trivedi M, et al.
Treatment of Depression: Newer Pharmacotherapies.
Rockville, Md: Agency for Health Care Policy and Research; 1999.

[37.](#)
Linde K, Mulrow CD.
St. John's Wort for depression [Cochrane Review on CD-ROM].
Oxford, England: Cochrane Library, Update Software; 1998:issue 4.

[38.](#)
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.
[MEDLINE](#)

[39.](#)
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.
[MEDLINE](#)

[40.](#)
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.

[41.](#)
American Psychiatric Association.
Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition.
Washington, DC: American Psychiatric Association; 1994.

[42.](#)
Media Metrix, Inc.
Internet ratings June 2000.
Available at: http://us.mediametrix.com/home.jsp. Accessed September 5, 2000.

[43.](#)
Internet.com Corp.
Search Engine Watch.

Search engine alliances chart. Available at:
http://www.searchenginewatch.com. Accessed June 10, 2000.

44.
Internet.com Corp.
Search Engine Watch.
How search engines rank Web pages. Available at:
http://www.searchenginewatch.com. Accessed June 10, 2000.

45.
Cyber Dialogue.
The number of health Web sites: plenty of options, relatively few choices.
Cybercitizen Health trend report No. 17. 2000. Available at:
http://www.cyberdialogue.com. Accessed December 10, 2000.

46.
PC Data.
Top twenty healthcare sites for September 2000. Who's on first? WebMD
moves up in rankings.
September 2000. Available at: http://www.pcdata.com. Accessed October
12, 2000.

47.
Carletta J.
Assessing agreement on classification tasks: the kappa statistic.
Computational Linguistics.
1996;22:249-254.

48.
Fry E.
The readability graph validated at primary levels.
The Reading Teacher.
1969;22:534-538.

49.
Fry E.
Fry's readability graph: clarifications, validity, and extension to level 17.
J Reading.
1977;21:242-252.

50.
Gilliam B, Pena S, Moutain L.
The Fry graph applied to Spanish readability.
The Reading Teacher.
January 1980:426-430.

51.
Kirsch IS, Jungeblut A, Jenkins L, Kolstad A.

Adult Literacy in America.
Washington, DC: National Center for Education Statistics; 1993.

52.
Kincaid JP, Fishburne RP, Rogers RL, Chissom BS.
Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel.
Memphis, Tenn: Naval Air Station; 1975.

AMA | INFO CENTERS   SHORT CUT: Go

Table 1

JAMA

Table

Please close this browser window to return to the article.
For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.

Health Information on the Internet

(*JAMA.* 2001;285:2612-2621)

Table 1. Selected English- and Spanish-Language Web Sites*

	URL Address	Conditions Examined
English-Language Web Sites		
Popular general health†		
Allhealth.com	www.allhealth.com	Breast cancer, childhood asthma, depression, obesity
CBS Health Watch	www.cbshealthwatch.com	Breast cancer, childhood asthma, depression, obesity
DrKoop.com	www.drkoop.com	Breast cancer, childhood asthma, depression, obesity
Intelihealth	www.intelihealth.com	Breast cancer, childhood asthma, depression, obesity
Onhealth	www.onhealth.com	Breast cancer, childhood asthma, depression, obesity
WebMD	www.webmd.com	Breast cancer, childhood asthma, depression, obesity
Condition-specific‡		
American Academy of Allergy, Asthma, and Immunology	www.aaaai.org	Childhood asthma
American Cancer Society	www.cancer.org	Breast cancer

Table 1

American Cancer Society	www.cancer.org	Breast cancer
American Obesity Association	www.obesity.org	Obesity
Athealth.com	www.athealth.com	Depression
Cancernet	www.cancernet.gov	Breast cancer
Depression.com§	www.depression.com	Depression
MyAsthma	www.myasthma.com	Childhood asthma
National Heart, Lung, and Blood Institute	www.nhlbi.nih.gov	Childhood asthma
National Institute of Mental Health	www.nimh.nih.gov	Depression
National Library of Medicine	www.nlm.nih.gov	Obesity
Obesity Online	www.obesity-online.com	Obesity
Oncolink	www.oncolink.com	Breast cancer
Search engine‡		
Yahoo	www.yahoo.com	Breast cancer, depression, obesity

Spanish-Language Web Sites

General health‡		
Graciasdoctor	www.graciasdoctor.com	Breast cancer, childhood asthma, depression, obesity
Salud	www.salud.com	Breast cancer, childhood asthma, depression, obesity
Salud Latina	www.saludlatina.com	Breast cancer, childhood asthma, depression, obesity
Condition-specific‡		
Cancernet	www.cancernet.gov	Breast cancer
Centro Peso	www.centropeso.com	Obesity
National Institutes of Health	www.nih.gov	Childhood asthma
New York Online Access to Health	www.noah-health.org	Depression

*URL indicates uniform resource locator.
†Top 6 sites ranked by Cyber Dialogue and PC Data Online, September 2000.

Table 1

‡Selected by project staff.
§This Web site was no longer accessible by *JAMA* and RAND staff as of May 7, 2001. Copies of the Web site content
  are available from the authors on request.

Please close this browser window to return to the article.

Figure 1

JAMA

Figure

Please close this browser window to return to the article.
For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.

Health Information on the Internet
(*JAMA.* 2001;285:2612-2621)

Figure 1. Flow of Search Engine Study

Please close this browser window to return to the article.

Table 2

JAMA

Table

Please close this browser window to return to the article.
For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.
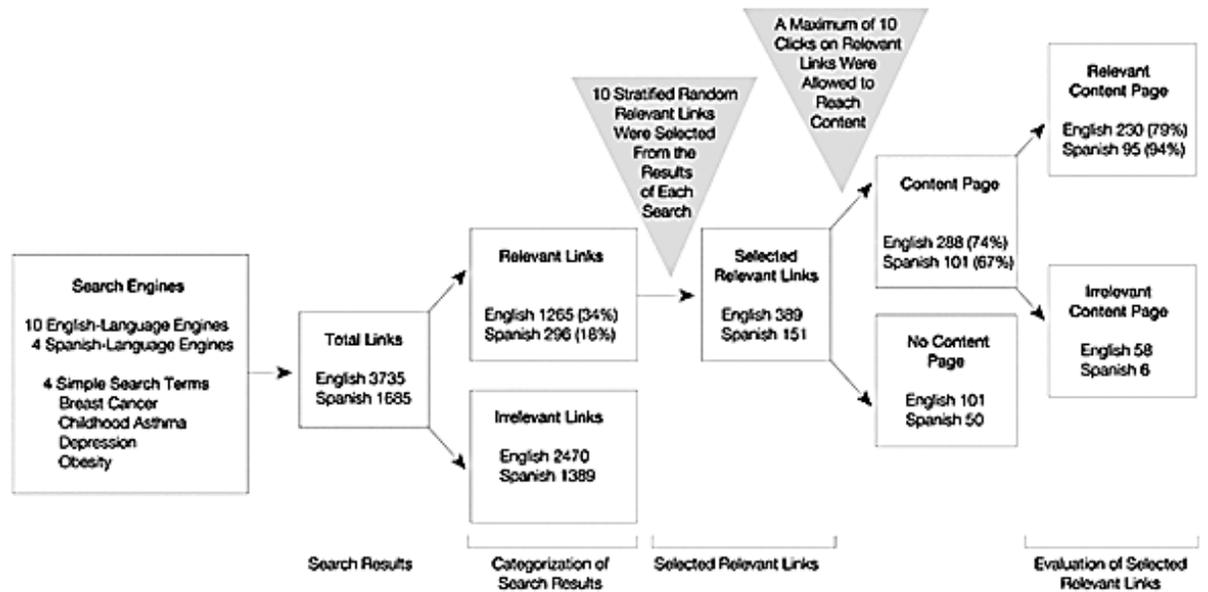
Health Information on the Internet

Table 2

JAMA

Table

Please close this browser window to return to the article.
For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.

Health Information on the Internet

(*JAMA.* 2001;285:2612-2621)

Table 2

Table 2. Search Engine Characteristics and Efficiency in Reaching Relevant Content

Search Engine	No. of Links on First Results Page (4 Medical Conditions)	Relevant Links From First Results Page, No. (%)	Selected Relevant Links Leading to Relevant Content, %	Proportion of All Links Leading to Relevant Content, %*
English-Language Engines				
Overall	3735	1265 (34)	59	20
Altavista	282	142 (50)	35†	18
Ask Jeeves	1198	299 (25)†	68	17
Direct Hit	302	129 (43)‡	38†	16
Excite	232	61 (26)†	60	16
Google	163	51 (31)	74‡	23
Goto	344	155 (45)‡	43†	19
Lycos	353	132 (37)	50	19
Metacrawler	384	153 (40)‡	70	28
Northern Light	241	55 (23)†	88‡	20
Yahoo	236	88 (37)	67	25
Spanish-Language Engines				
Overall	1685	296 (18)	63	12
Quepasa	283	52 (18)	59	11
Te Respondo	853	105 (12)†	78‡	9
Yahoo Espanol	181	63 (35)‡	66	23
Yupi	368	76 (21)	49†	10

*Assuming randomly selected relevant links are representative of all relevant links.
†Statistically worse performance than the mean of other search engines of this language ($P \leq .05$).
‡Statistically better performance than the mean of other search engines of this language ($P \leq .05$).

Please close this browser window to return to the article.

Table 2

Table 2

Table 2. Search Engine Characteristics and Efficiency in Reaching Relevant Content

Search Engine	No. of Links on First Results Page (4 Medical Conditions)	Relevant Links From First Results Page, No. (%)	Selected Relevant Links Leading to Relevant Content, %	Proportion of All Links Leading to Relevant Content, %*
English-Language Engines				
Overall	3735	1265 (34)	59	20
Altavista	282	142 (50)	35†	18
Ask Jeeves	1198	299 (25)†	68	17
Direct Hit	302	129 (43)‡	38†	16
Excite	232	61 (26)†	60	16
Google	163	51 (31)	74‡	23
Goto	344	155 (45)‡	43†	19
Lycos	353	132 (37)	50	19
Metacrawler	384	153 (40)‡	70	28
Northern Light	241	55 (23)†	88‡	20
Yahoo	236	88 (37)	67	25
Spanish-Language Engines				
Overall	1685	296 (18)	63	12
Quepasa	283	52 (18)	59	11
Te Respondo	853	105 (12)†	78‡	9
Yahoo Espanol	181	63 (35)‡	66	23
Yupi	368	76 (21)	49†	10

*Assuming randomly selected relevant links are representative of all relevant links.
†Statistically worse performance than the mean of other search engines of this language ($P \leq .05$).
‡Statistically better performance than the mean of other search engines of this language ($P \leq .05$).

Please close this browser window to return to the article.

Table 2

JAMA

Index of Figures and Tables

TABLES

▼

▲

FIGURES

▼

Please close this browser window to return to the article.

Health Information on the Internet
(*JAMA.* 2001;285:2612-2621)

FIGURES
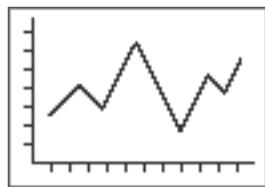
Figure 1. Flow of Search Engine Study

Figure 2. Reading Grade Levels for Selected Web Sites by Language

TABLES

Table 1. Selected English- and Spanish-Language Web Sites*

Table 2. Search Engine Characteristics and Efficiency in Reaching Relevant Content

Table 3. Comparison of Coverage and Accuracy of Selected Condition-Related Topics for 4 Conditions

Table 4. Presence of Conflicting Information on English-Language Web Sites

Please close this browser window to return to the article.

© 2001 American Medical Association. All rights reserved.

Figure 2

JAMA

Figure

Please close this browser window to return to the article.
For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.

Health Information on the Internet
(*JAMA.* 2001;285:2612-2621)

Figure 2. Reading Grade Levels for Selected Web Sites by Language

Figure 2

Reading grade levels were determined using the Fry Readability Graph method.[48-50]

Please close this browser window to return to the article.

Please close this browser window to return to the article.

Table 3

Please close this browser window to return to the article.
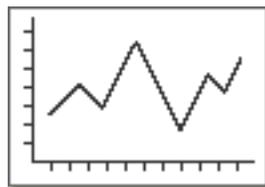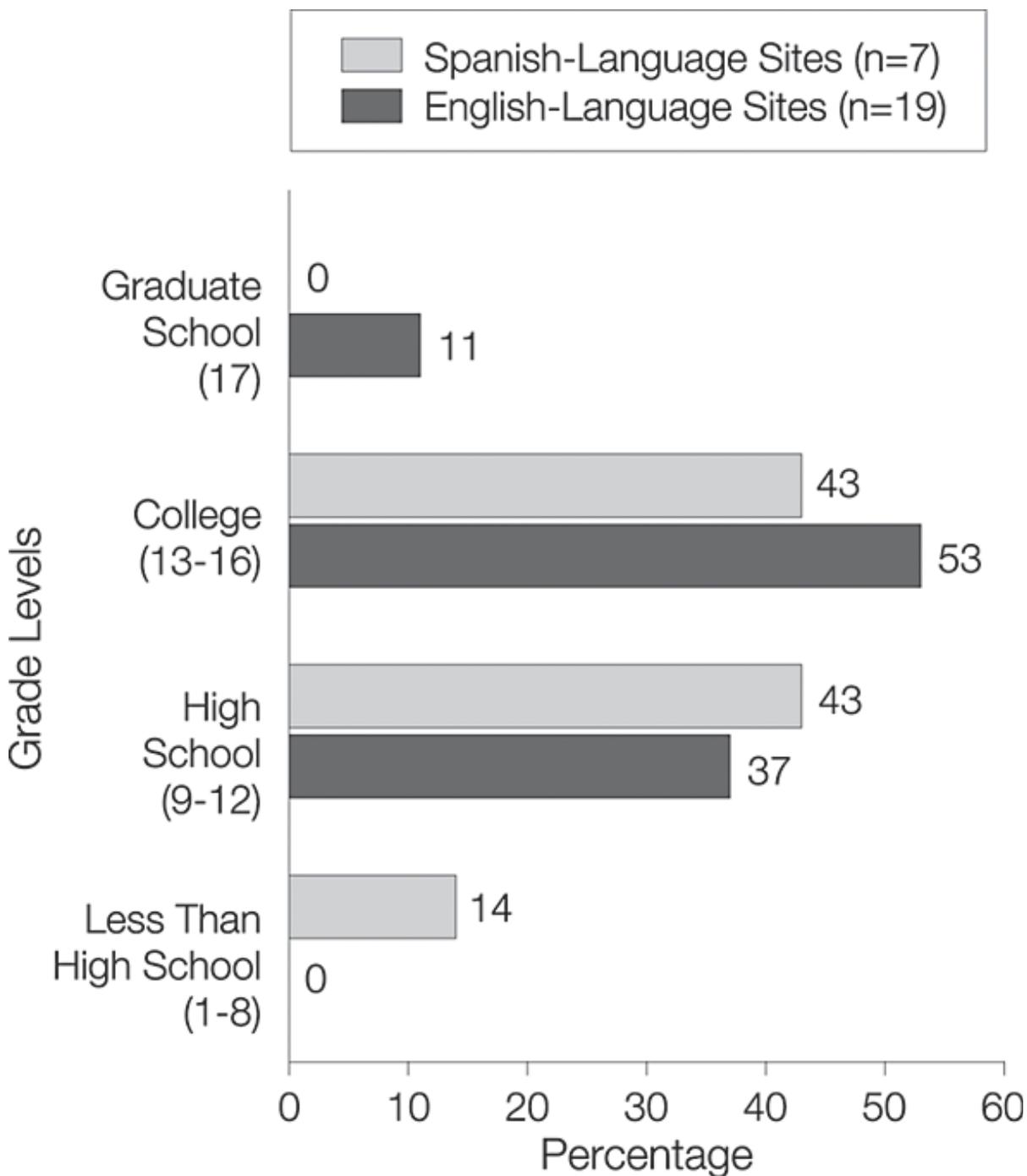For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.

Health Information on the Internet

(*JAMA.* 2001;285:2612-2621)

Table 3. Comparison of Coverage and Accuracy of Selected Condition-Related Topics for 4 Conditions

Condition-Related Topics	Average for 10 English-Language Web Sites per Condition*							Average for 4 Spanish-Language Web Sites per Condition						
	Coverage, %			Correctness, %			More Than Minimal Coverage and Completely Correct, %	Coverage, %			Correctness, %			More Than Minimal Coverage and Completely Correct, %
	None	Minimal	More Than Minimal	Mostly Not	Mostly	Completely		None	Minimal	More Than Minimal	Mostly Not	Mostly	Completely	
Breast Cancer														
Overall§	16†	17	67†	0	9	91	63†	49	12	39†	0	4	96	39†
Risk assessment and use of tamoxifen for risk reduction	10	12	78	0	11	89	73	45	14	41	0	17	83	36
Screening	10	11	79	1	13	86	69	27	12	61	0	0	100	61
Evaluation of a palpable breast mass	18	25	57	0	7	93	57	30	20	50	0	3	97	50
Treatment	13	14	73	2	6	92	70	61	12	27	0	12	88	27
Alternatives to standard surgical and medical treatments	28	23	49	0	4	96	49	90	0	10	0	0	100	10
Childhood Asthma														
Overall§	27	30	43‡	3	13	84	36‡	33†	40	27	4	43	53	23
Symptoms	33	26	41	2	9	89	36	25	31	44	3	36	61	40
Poorly controlled asthma	48	29	23	19	9	72	18	42	33	25	14	29	57	25
Therapies and adverse effects	13	22	65	1	23	76	48	27	44	29	6	38	56	27
Initial management of severe asthma	33	46	21	1	15	84	19	40	50	10	3	59	38	10
Risk factors	29	32	39	0	16	84	32	32	42	26	0	53	47	21
Etiology	32	22	46	0	2	98	46	31	36	33	4	44	52	23
Expectations from therapy	23	41	36	0	10	90	36	44	39	17	0	44	56	14
Depression														

Table 3

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Overall§	20†	27	53	3	22	75	44	61	24	15‡	18	19	63	12‡
Symptoms	13	15	72	1	17	82	61	45	37	18	5	31	64	14
Treatments	17	11	72	2	30	68	56	42	22	36	9	36	55	24
Antidepressant medications	16	17	67	0	22	78	55	77	14	9	0	6	94	9
Role of counseling	31	29	40	1	26	73	33	64	27	9	17	30	53	9
Suicidal ideation	12	46	42	5	11	84	37	55	40	5	14	36	50	5
Evaluation	33	54	13	6	23	71	8	84	16	0	67	0	33	0
Etiology	0	3	97	3	7	90	87	27	9	64	13	12	75	45
Obesity‖														
Overall§	35‡	25	40‡	3	11	86	37‡	69‡	15	16	0	32	68	15
Definitions and indications for weight loss	36	16	48	4	18	78	42	50	19	31	0	19	81	31
Health risks	12	29	59	1	9	90	56	23	50	27	0	48	52	21
Risk and benefits of popular diets	49	34	17	5	8	87	17	100	0	0	...	...	...	0
Physical activity/ prevention	20	34	46	0	6	94	43	54	17	29	0	37	63	29
Medications endorsed for weight loss	47	25	28	0	17	83	25	79	11	10	0	50	50	8
Surgery	32	22	46	8	7	85	44	90	6	4	0	29	71	4
Safety and effectiveness of dietary supplements	61	19	20	0	4	96	19	100	0	0	...	...	...	0
Mean of all 4 conditions	24	25	51	2	13	84	45	53	23	24	6	24	70	22

*Averages taken from 10 English-language Web sites with the exception of childhood asthma, for which 9 were used.
†Statistically better performance than condition average within the same language ($P \leq .05$).
‡Statistically lower performance than condition average within the same language ($P \leq .05$).
§Weighting all clinical elements equally, rather than weighting all topics equally. See Online Table 2.
‖Ellipses indicate that an element was not addressed and could not be assessed as correct.

Please close this browser window to return to the article.

Table 4

JAMA

Table

Please close this browser window to return to the article.
For enhanced viewing of figures and tables, it is recommended that the PDF version be downloaded.

Health Information on the Internet

(*JAMA.* 2001;285:2612-2621)

Table 4. Presence of Conflicting Information on English-Language Web Sites

	% of Reviews Noting Any Conflict	% of Reviews With Any Conflicting Information by Category					
		Treatment	Diagnosis	Definition	Adverse 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 fewer reviews for this condition noted conflicts of this type than for the mean condition ($P<.05$).
†Significantly more reviews for this condition noted conflicts of this type than for the mean condition ($P<.05$).

Please close this browser window to return to the article.