

# Empirical Studies Assessing the Quality of Health Information for Consumers on the World Wide Web

## A Systematic Review

Gunther Eysenbach, MD

John Powell, MSc, MRCPsych, MFPHM

Oliver Kuss, PhD

Eun-Ryoung Sa, MS

**T**HE INTERNET HAS BECOME AN important mass medium for consumers seeking health information and health care services online.<sup>1</sup> A recent concern and public health issue has been the quality of health information on the World Wide Web. However, the scale of the problem and the "epidemiology" (distribution and determinants) of poor health information on the Web are still unclear, as is their impact on public health and the question of whether poor health information on the Web is a problem at all.<sup>2</sup> Many studies have been conducted to describe, critically appraise, and analyze consumer health information on the Web. These typically report proportions of inaccurate or imperfect information as estimates of the prevalence of flawed information or the risk of encountering misinformation on the Web.

However, to date no systematic and comprehensive synthesis of the methodology and evidence has been attempted. Two previous systematic reviews focused on compiling quality criteria and rating instruments, but did not synthesize evaluation results. Jadad and Gagliari<sup>3</sup> reviewed non-research-based rating systems (eg, cri-

**Context** The quality of consumer health information on the World Wide Web is an important issue for medicine, but to date no systematic and comprehensive synthesis of the methods and evidence has been performed.

**Objectives** To establish a methodological framework on how quality on the Web is evaluated in practice, to determine the heterogeneity of the results and conclusions, and to compare the methodological rigor of these studies, to determine to what extent the conclusions depend on the methodology used, and to suggest future directions for research.

**Data Sources** We searched MEDLINE and PREMEDLINE (1966 through September 2001), Science Citation Index (1997 through September 2001), Social Sciences Citation Index (1997 through September 2001), Arts and Humanities Citation Index (1997 through September 2001), LISA (1969 through July 2001), CINAHL (1982 through July 2001), PsychINFO (1988 through September 2001), EMBASE (1988 through June 2001), and SIGLE (1980 through June 2001). We also conducted hand searches, general Internet searches, and a personal bibliographic database search.

**Study Selection** We included published and unpublished empirical studies in any language in which investigators searched the Web systematically for specific health information, evaluated the quality of Web sites or pages, and reported quantitative results. We screened 7830 citations and retrieved 170 potentially eligible full articles. A total of 79 distinct studies met the inclusion criteria, evaluating 5941 health Web sites and 1329 Web pages, and reporting 408 evaluation results for 86 different quality criteria.

**Data Extraction** Two reviewers independently extracted study characteristics, medical domains, search strategies used, methods and criteria of quality assessment, results (percentage of sites or pages rated as inadequate pertaining to a quality criterion), and quality and rigor of study methods and reporting.

**Data Synthesis** Most frequently used quality criteria used include accuracy, completeness, readability, design, disclosures, and references provided. Fifty-five studies (70%) concluded that quality is a problem on the Web, 17 (22%) remained neutral, and 7 studies (9%) came to a positive conclusion. Positive studies scored significantly lower in search ( $P=.02$ ) and evaluation ( $P=.04$ ) methods.

**Conclusions** Due to differences in study methods and rigor, quality criteria, study population, and topic chosen, study results and conclusions on health-related Web sites vary widely. Operational definitions of quality criteria are needed.

JAMA. 2002;287:2691-2700

www.jama.com

**Author Affiliations:** Unit for Cybermedicine and E-Health, Department of Clinical Social Medicine, University of Heidelberg, Heidelberg, Germany (Dr Eysenbach); Health Services Research Unit, London School of Hygiene and Tropical Medicine, London, England (Dr Powell); Department of Medical Epidemiology, Biometry and Informatics, University of Halle-Wittenberg, Halle/Saale, Germany (Dr Kuss); and Global Health

Network Group, Department of Epidemiology, University of Pittsburgh, Pittsburgh, Pa (Ms Sa). Dr Eysenbach is now with the Centre for Global eHealth Innovation, Toronto General Hospital, Toronto, Ontario.

**Corresponding Author and Reprints:** Gunther Eysenbach, MD, Centre for Global eHealth Innovation, Toronto General Hospital, 190 Elizabeth St, Toronto, Ontario, Canada M5G 2C4 (e-mail: ey@yi.com).

**For editorial comment see p 2713.**

teria used for “awards,” “seals of approval,” or gateways) published on the Internet but did not include rating systems published in the peer-reviewed literature. Kim et al<sup>4</sup> compiled instruments or criteria proposed to evaluate Web sites, but it is unclear to what extent and with what results these criteria can be and have been applied in practice.

To our knowledge, this is the first systematic review conducted to compile criteria actually used and to synthesize evaluation results from studies containing quantitative data on structure and process measures of the quality of health information on the Web. The objectives of this study were to establish a methodological framework on how “quality” on the Web is evaluated in practice, to determine the heterogeneity of the results and conclusions, to compare the methodological rigor of these studies, to determine to what extent the conclusions depend on the methodology used, and to suggest future directions for research.

## METHODS

### Inclusion and Exclusion Criteria

Inclusion and exclusion criteria were established in advance in a written protocol. Studies were included when their authors: (1) searched the World Wide Web systematically for health information (eg, to identify all consumer Web sites for a given topic) or clearly defined a set of specific health information services to be included; (2) evaluated information likely to be accessed by consumers; (3) evaluated the quality of health information or services against certain criteria—for example, by judging the authority of source; assessing the accuracy of information, readability, or comprehensiveness; or by comparing Internet information and services against those outside of the Internet; and (4) provided quantitative results, such as proportions of Web sites complying with the quality criteria, or the distribution of quality scores, or information on the readability level(s) in the sample.

Studies were excluded if authors simply listed “quality” Web sites without in-

dicating that they had performed a comprehensive, systematic search, if they did not list quality criteria, or if they did not quantitate how many Web sites were assessed in total or how many did not comply with quality criteria. We also excluded studies assessing tools that enable or facilitate access to health information, such as search engines, directories, health portals, rating systems, and review services, or studies comparing different methods of information retrieval. Theoretical papers that only described quality criteria alone were also excluded, as were qualitative descriptions of Internet contents. We did not include studies that dealt with content other than health information, or studies evaluating information not intended for consumers (eg, those providing continuing medical education), but we did include studies evaluating services and information likely to be used by both professionals and consumers (eg, those providing drug information). Studies evaluating a single site or application were excluded, as were studies focusing exclusively on privacy and security issues, or on interactions via newsgroups or e-mail.

### Search Strategy

We sought all relevant studies and unpublished reports, regardless of language or peer-review/publication status. We searched MEDLINE and PREMEDLINE (1966-September 2001) by entering the following query into the PubMed interface on September 20, 2001 (*quality OR reliability OR accuracy OR readability OR evaluation OR assessment*) AND (*information OR education OR advice*) AND (*internet OR web OR ehealth OR “e-health” OR cyber\* OR www*). We retrieved 1545 citations and their titles and abstracts were screened for potential relevance by 2 independent reviewers. In addition, 1 reviewer used analogous search terms to screen citations in LISA (Library and Information Science Abstracts Database, 1969- July 2001) (1269 hits), CINAHL Nursing and Allied Health (1982-July 2001) (2312 hits), PsychINFO (American Psychological Association, 1988-September 2001 [week

2]) (321 hits), EMBASE (Elsevier Science BV, Amsterdam/NL, 1988-June 2001) (647 hits), and SIGLE (Grey Literature in Europe database, 1980-June 2001) (83 hits). The Web of Science database (Institute for Scientific Information, Philadelphia, Pa), which includes the Science Citation Index (1997-September 2001), the Social Sciences Citation Index (1997-September 2001), and the Arts and Humanities Citation Index (1997-September 2001) was searched using both a traditional backward search strategy (1120 hits) as well as using a forward citation-search strategy, using 5 seminal papers in this field<sup>3-7</sup> as “seed publications” to identify all publications that subsequently cited 1 of these papers (533 hits). We also checked the references of all identified studies, hand-searched the *Journal of Medical Internet Research* (because it contains pertinent studies and because at that time it was not yet indexed on MEDLINE), conducted general Internet searches using combinations of the search terms at the Google and Northern Light search engines, reviewed 3 bibliographies of consumer health information studies collated by academic institutions, and used private literature databases.

In total, we screened 7830 citations and retrieved 165 potentially eligible full articles to determine whether they met the inclusion criteria. Articles in Hungarian, Japanese, French, Spanish, Dutch, and Italian were translated by professional translators into English or German.

### Data Extraction

Two reviewers (G.E., J.P.) independently extracted study characteristics using electronic data extraction forms, maintaining all data in an Access database (Microsoft Corp, Redmond, Wash). Each extraction form contained 85 items, of which 43 were closed questions, mostly to be answered on a scale (yes/no/partially/not applicable/to be discussed). We extracted information on search strategies used, quality criteria applied, and methodology of quality assessment used by authors, as well as quality and rigor of

study methodology and reporting (see below and TABLE 1).

We rated the overall tone of the authors' conclusions as "negative" if authors were wary and pessimistic about the Web as a source for health information, "positive" if authors did not express concern or recommended the Web as a source for health information, or "neutral" if authors discussed both risks and benefits, or if the reviewers disagreed.

To extract evaluation results of each study, one reviewer tabulated data on how many Web sites or Web pages were evaluated for each quality criterion and how many of those evaluated did not comply with the respective criterion.

### Statistical Analysis

To determine interobserver reliability for all coded items, we calculated standard  $\kappa$  values for each extracted item reported on a binary scale (eg, include study: yes/no), or weighted  $\kappa$  values for ordinal outcomes (eg, study characteristics graded on a scale [yes/partially/no]). We decided in advance not to include items with  $\kappa$  values less than 0.6.<sup>8</sup> For the final consensus coding used for analysis, reviewers examined all disagreements or items coded "to be discussed" and resolved them by extensive discussion.

We used meta-analysis of proportions<sup>9</sup> to test for homogeneity of results for quality criteria that have been used by 3 or more authors to assess Web sites. We pooled study results for a given quality criterion if the assumption of homogeneity could not be rejected on a significance level of less than .05.

To test whether the different conclusions of studies could be explained by their different quality scores we compared study assessment scores of study groups using the Wilcoxon test and logistic regression. All calculations were performed using SAS versions 6.12 and 8.2 (SAS Institute Inc, Cary, NC).

### Quality Scores

We summarized the quality of reporting and the rigor of the studies by checking whether authors reported

certain items deemed important for a well-reported, systematic, and comprehensive Internet search strategy (ie, S-score for search quality; range, 0-8 [transformed to percentage score]), and items thought to indicate a rigorous evaluation (ie, E-score for evaluation quality; [range, 0-10 [transformed to percentage score]]). To ascertain face validity, an initial list of candidate items was compiled by 2 reviewers (G.E., J.P.) independently and later combined by consensus. As studies

systematically evaluating Web information resemble in some ways systematic research overviews, the candidate criteria for the S-score partly derived from corresponding checklists.<sup>10</sup> To ascertain score reliability, we eliminated items for which  $\kappa$  was less than 0.6; the remaining items are shown in Table 1. One point was given per item reported in order to calculate a raw summary score for search quality and evaluation quality. The final reported S-score and E-score

**Table 1.** Summary of Characteristics of Included Studies, as Extracted by 2 Reviewers\*

Study Characteristics	Yes	No	Partially	NA	$\kappa$
Total included studies	79	85	...	...	0.93
Assessment target					
Accuracy	47	30	2†	...	0.98
Completeness	19	39	21‡	...	0.90
Technical criteria	53	26	...	...	0.77
Readability	11	63	5§	...	0.89
Design (aesthetics)	15	64	...	...	0.89
Search quality (S-score)					
Search date/period mentioned	51	26	2	...	0.96
Search tools mentioned	60	13	4	2	0.77
More than 1 search tool used	54	21	2¶	2	0.92
Search terms mentioned	54	16	1	8#	0.95
Consumer involvement, eg, when devising search strategy	3	75	1**	...	0.85
Initial hits reported	25	49	1	4	0.83
Sites in more than 1 language assessed	8	71	...	...	0.86
Interrater reliability for site selection determined	5	67	...	7	0.71
Evaluation quality (E-score)					
Raters blinded for the source	2	74	...	3††	1.00
Number of raters reported	39	37	...	3††	0.71
More than 1 rater	21	45	10‡‡	3††	0.83
Interrater reliability figure for evaluation determined	16	57	4	2	0.79
A priori criteria defined for accuracy	17	30	2	30	0.75
A priori criteria defined for completeness	24	16	...	39	0.87
Criterion standard for accuracy stated and different from personal opinion	32	17	...	30	0.97
Criterion standard for accuracy derived from clinical guideline or systematic review	15	34	...	30	1.00
Criterion standard for completeness stated and different from personal opinion	17	23	...	39	0.72
Criterion standard for completeness derived from clinical guideline or systematic review	10	30	...	39	0.93

\*Disagreements between reviewers were resolved by discussion.  $\kappa$  Values are reported as an indication for extraction reliability. NA indicates not applicable; ellipses, no occurrence.

†Authors claimed to have evaluated "value and amount of text" without specifying how "value" was defined.<sup>36,90</sup>

‡Completeness evaluated as part of accuracy.

§Did not use readability formulas, but subjectively assessed readability or understandability of content.

||Defined the search date very broadly, eg, just mentioned the year.

¶Used 1 meta-search engine, and nothing else.

#Studies that used only catalogs (eg, Yahoo) as search tools were rated NA if search terms were not mentioned.

\*\*Patient advocates involved in devising questions.

††Studies did not use human raters, as they used automatic methods only to determine reading level.

‡‡A subset of Web sites was evaluated by more than 1 rater, a second person was consulted when in doubt, or the rating instrument was pilot-tested by 2 raters to determine interrater reliability.

in ONLINE TABLE A (<http://jama.com>) represent the percentages of the maximum score achievable for each study (because some items are not applicable for some studies).

To test construct validity we compared the scores of 5 letters to the editor with the scores of the 74 full publications. As expected, the letters had, on average, lower S-scores (28.8 vs 44.2,  $P=.06$ ), but similar E-scores (22.4 vs 33.5,  $P=.45$ ). The 7 full articles published in peer-reviewed non-English-language journals had lower average E-scores than articles published in English-language journals (mean E-score: 18.6 vs 34.8,  $P=.09$ ), while their S-scores were similar (mean S-score: 42.3 vs 44.5,  $P=.85$ ). While these differences are not

significant at a 5% level, perhaps due to the small study number, the trend goes in the expected direction.

## RESULTS

### Included Studies

Of 170 articles reviewed, 85 met the inclusion criteria<sup>11-95</sup> (Online Table A). These constitute 79 distinct studies, as 6 studies were published either in duplicate,<sup>26,27,50,51,92,93</sup> in a fragmented manner,<sup>87,88</sup> or were reported both in a peer-reviewed publication and in an unpublished form.<sup>14,15,68,69</sup> In these cases we pooled the results from the different articles and referred to these as a single study. Table 1 summarizes, and Online Table A lists, all included studies.

### Quality Criteria and Methods Used to Evaluate Web Sites

**Technical Criteria (T).** "Technical" quality criteria were defined as general, domain-independent criteria, ie, criteria referring to the question of how the information was presented or what meta-information was provided. The 24 technical quality criteria most frequently used (TABLE 2) are variations of what could be called "transparency criteria" from the print world, mentioned by Silberg et al<sup>7</sup>: authorship, attribution, disclosure, and currency. While the latter is strictly speaking a dimension of accuracy, almost all studies sought for provision of a date of creation or last update (rather than actual currency of the content), which is a technical criterion.

**Design (D).** Fifteen studies also evaluated subjective design features such as the visual aspect of the site or layout, but only a few studies reported results (ONLINE TABLE B), presumably because of their subjectivity and low reliability. For example, Gillois et al<sup>39</sup> used a visual analog scale to assess the quality of visual presentation and of the interface, but noted high interobserver variability. Stausberg et al<sup>81</sup> report a  $\kappa$  of 0.08 when evaluating navigation, and 0.23 when evaluating layout. Speed, browser compatibility, and presence of a search engine were considered technical criteria.

**Readability (R).** Eleven studies used readability formulas to establish the reading level of a document based on the complexity and length of words and sentences (Online Table B). Nine studies used the Flesch-Kincaid (FK) Grade Level Index. Other formulas used include the SMOG Readability Formula, the Fry Readability Graph, the Gunning-Fog formula, and the Lexile Framework. As different formulas yield different reading levels for the same document, and as authors use different cutoffs, the results cannot be pooled, but studies suggest that reading levels are frequently too high (Online Table B).<sup>65</sup>

Using reading formulas has limitations, as readability scores do not reflect other factors that affect comprehension such as frequency and explanation of medical jargon, writing style

**Table 2.** Overview of Quality Criteria (Other Than Accuracy, Completeness, Readability Level, and Design) Used by 3 or More Studies and Pooled Evaluation Results\*

Quality Criterion	No. of Studies	Total No. of Web Sites/Web Pages Evaluated	Sites/Pages Not Complying, %†	P Value
(a) Disclosure of authorship	19	1636	(57.8)‡	<.001
(b) Disclosure of ownership	5	196	1.0	.28
(c) Sources clear	4	110	(44.5)‡	<.001
(d) Disclosure of sponsorship	7	738	(93.4)§	.006
(e) Disclosure of advertising	3	119	30.3	.11
(f) Statement of purpose	4	230	(51.7)‡	<.001
(g) General disclosures	3	298	(49.7)‡	.001
(h) Date of creation disclosed	5	284	83.1	.56
(i) Date of last update disclosed	7	801	(58.8)§	.049
(j) Date of creation or update disclosed	12	1366	(63.7)‡	<.001
(k) Authors' credentials disclosed	9	1030	(70.6)‡	<.001
(l) Credentials of physicians disclosed	3	81	97.5	.90
(m) Authors' affiliation disclosed	5	779	(42.5)‡	<.001
(n) Easy navigation (subjective rating)	4	326	(10.1)‡	<.001
(o) Internal search engine present	3	91	79.1	.16
(p) Links provided	4	238	(27.7)§	.005
(q) References provided	30	2135	(68.9)‡	<.001
(r) Balanced evidence	3	182	(47.9)‡	<.001
(s) Writing style appropriate (subjective rating)	4	136	16.2	.84
(t) Feedback mechanisms provided	4	157	(14.0)‡	<.001
(u) Fax number provided	5	1322	(38.0)§	.02
(v) E-mail address provided	8	1642	(24.3)‡	.001
(w) General disclaimers provided	6	390	(75.9)§	.047
(x) Copyright notice	4	318	40.9	.09
(y) Editorial review process	5	166	87.3	.75
(z) Hierarchy of evidence clear	4	89	76.4	.75

\*Most criteria classified as "technical" criteria, except "balanced evidence" (coded as "completeness") and "writing style appropriate" (coded as "readability").

†Proportions in parentheses must be interpreted with caution, because these are based on studies whose results vary significantly (ie, are statistically heterogeneous).

‡Highly significant heterogeneity ( $P \leq .001$ ); results not sufficiently homogeneous to be pooled.

§Significant heterogeneity ( $P \leq .05$ ); results not sufficiently homogeneous to be pooled.



(use of active voice, nonpatronizing language, motivational messages, tone/mood, how it relates to the audience), or use of culturally specific information. Few studies analyzed these important but subjective aspects—Oermann and Wilson<sup>66</sup> discussed some of these parameters, Fitzmaurice and Adams<sup>35</sup> scored the writing style, and Wilson et al<sup>94</sup> examined the cultural sensitivity of the documents to various ethnic groups using Bloch's Ethnic/Cultural Assessment Tool. None of the studies conducted comprehension tests with actual consumers or used judgments of literacy experts.

**Accuracy (A).** Accuracy can be defined as the degree of concordance of the information provided with the best evidence or with generally accepted medical practice. Alternative terms used for the concept of "accuracy" include "reliability"<sup>51</sup> or "conventionality of information."<sup>79</sup> We coded 2 studies as assessing accuracy "partially," because authors claimed to have evaluated "value and amount of text" without further specifying how they define "value."<sup>36,90</sup>

Authors who appraised the accuracy and/or comprehensiveness could either define clear criteria beforehand (a priori) or extract information from the Web first and then check these claims against the literature (a posteriori). Thirty studies assessed accuracy a posteriori, and only 19 a priori (Online Table A). Studies defining elements a priori varied considerably in the granularity and specificity of the items defined in advance: for example, 1 study stated that "paracetamol should be given in a dose of 10 to 15 mg/kg every 4 hours,"<sup>51</sup> while others just predefined broad keywords that should be mentioned, eg, "oestrogen."<sup>75</sup>

Accuracy ideally should be defined using the best available evidence.<sup>5</sup> Fourteen studies used evidence-based guidelines or systematic reviews to define elements a priori and 1 study used a guideline as the criterion standard a posteriori. Eleven studies used primary literature for a posteriori comparison; however, with 1 exception<sup>89</sup> these studies did not provide details on how and what lit-

erature was screened, leaving open the possibility that some authors did not actually search the literature, but just compared information against their own knowledge of the literature. Textbooks or expert consensus were used as the criterion standard in 3 a priori and in 3 a posteriori studies.<sup>29,54,55,64,70,84</sup> The remaining studies used unclear sources or the personal opinion of the author as the a priori (2 studies) or a posteriori (15 studies) criterion standard.

**Completeness/Comprehensiveness/Coverage/Scope (C).** Several methods were used to evaluate "completeness," "comprehensiveness," "coverage," or "scope." Most authors<sup>8</sup> calculated a proportion of a priori-defined elements covered by a Web site or reported the proportion of Web sites that mentioned all key elements (eg, from a clinical guideline). Willems and Bouvy<sup>92,93</sup> used a 5-point scale to evaluate completeness. One study<sup>11</sup> evaluated "balance," eg, whether the adverse effects and contraindications as well as the advantages of a drug are presented, but detailed subcriteria were not made explicit. The Soot score<sup>13,46,58,79</sup> addresses completeness by measuring coverage of topic areas defined a priori, weighting some areas as being more important than others.

While 19 studies evaluated completeness as a distinct entity, 21 other studies evaluated completeness as an integral part of accuracy (marked "AC" in Online Table A). For example, Impicciatore et al<sup>51</sup> reported that only 4 of 44 Web sites "adhered closely to the guidelines," implying that these sites were both accurate and complete, whereas Davison<sup>27</sup> only assessed the accuracy of statements made on the site against the respective statement in the guideline without requiring that all recommendations from nutritional guidelines appear on the Web site. While it is possible to evaluate accuracy without demanding completeness, studies in which authors just prepared checklists for completeness sometimes raise the question as to how and to what extent

accuracy has been evaluated.<sup>29,33,54,55,75</sup> For example, the Soot<sup>79</sup> score is a weighted checklist giving points for each broad topic covered by the site (such as "treatment options"), but does not specify whether topics were merely present or absent or if they were evaluated for accuracy.

Of the 19 studies that evaluated completeness as a distinct entity, 5 studies used an external source a priori to define elements: 2 used evidence-based guidelines, 2 used primary literature, and 1 used a textbook. The remaining studies used unclear sources or the personal opinion of the author, either a priori (7 studies) or a posteriori (7 studies).

### Score Systems

Twenty-two studies<sup>†</sup> used a composite score system, most frequently the Soot score,<sup>13,46,58,79</sup> for adherence to quality criteria. Two studies<sup>38,87</sup> used DISCERN<sup>96</sup> and another used Suitability Assessment of Materials (SAM),<sup>65</sup> both described by their developers as "validated" instruments to assess printed patient education material; however, none of these have been validated in a way such that a higher score would predict better health outcomes or consumer satisfaction.

### Quality of Studies

The quality scores for each study are shown in Online Table A. Out of the 79 studies, 49 received an S-score and an E-score of 50 or less, indicating that both search strategy and evaluation were not well reported or rigorous. However, of the 31 studies using more than 1 rater, 20 made efforts to ascertain the within-study reliability of their instruments, which is in contrast to what has been criticized in a previous study about Web-based instruments.<sup>97</sup>

### Study Conclusions and Study Quality

Most studies (55 [70%]) concluded that quality is a problem on the Internet ( $\kappa$

\*References 12-15, 22, 29, 35, 37, 42, 44-46, 50, 51, 54-56, 58, 59, 70, 75, 78, 79, 85, 92, 93, 95.

†References 11, 13-15, 23, 31, 35, 36, 38, 42-44, 46, 58, 59, 70, 72, 74, 75, 78, 79, 87, 90.

for coding conclusions, 0.88). Seventeen (21.5%) were neutral. Only 7 (9%) came to a more positive conclusion, none of which used evidence-based guidelines as a criterion standard. The mean S-score of the positive studies was significantly lower than that of the negative studies (29.4 vs 45.0;  $P = .02$ ), as was the E-score (15.6 vs 37.9;  $P = .04$ ), indicating that the more enthusiastic studies used a less rigorous search and evaluation strategy than did negative studies. In a logistic regression model with both scores as predictive variables and the study conclusion "positive/negative" as the dependent variable, the odds ratio (OR) for S-score is 0.929 (95% confidence interval [CI], 0.871-0.991;  $P = .02$ ), ie, for each additional S-score percentage point the odds to reach a positive conclusion decreases by 7.1%; the OR for E-score is 0.938 (95% CI, 0.884-0.996;  $P = .04$ ), ie, for each additional E-score percentage point the odds for a positive conclusion decreases by 6.2%.

When studies evaluated accuracy, the proportions of inaccurate Web sites depended on the level of evidence used as a criterion standard: studies that didn't report the criterion standard or that used personal opinion found an average of 15.4% of Web sites inaccurate; those using literature, textbooks, or expert consensus, 35.3%; and those using clinical guidelines, 38.3%. In an ordinal logistic regression model, the reported proportions of inaccurate Web sites were significantly associated with the 3 levels of evidence used by authors as criterion standard ( $P < .001$ ).

Even if authors had similar results, they sometimes interpreted these differently. For example, 3 studies found that about 5% of cancer Web sites provided inaccurate information, but these results were interpreted as being "of concern,"<sup>16</sup> "encouraging,"<sup>76</sup> or even "reassuring."<sup>46</sup> In at least 1 case investigators reviewed the same topic area, but arrived at opposite conclusions.<sup>28,75</sup>

## COMMENT

We reviewed 79 studies in which authors evaluated a total of 5941 Web sites and 1329 Web pages, and reported 408

evaluation results for 86 distinct quality criteria (Online Table B).

## Content Quality

In our review, most authors who evaluated content found significant problems, criticizing lack of completeness, difficulty in finding high-quality sites, and lack of accuracy, in particular if "accuracy" also implied "completeness."

Five of eight studies reporting results on completeness found that around 90% of Web sites were "incomplete" (Online Table B). However, completeness as a requirement has questionable validity from the perspective of the user or the public health researcher. First, too much information may overburden users. Web sites may deliberately and with good reason focus on a single topic in-depth rather than aiming for comprehensiveness. Second, in contrast to printed educational material, a single Web page or Web site is part of a universe of information: a topic not covered by one Web page or site may be covered by another (perhaps linked) Web page. Consumers will usually search across different Web sites when looking for specific health information.<sup>98</sup> Mechanical comparison of elements from a guideline with elements covered by a single Web site without taking into account the context and purpose of the site or exploring links to other sites is of limited use. Perhaps a better approach would be to evaluate whether materials cover the topics they claim to be discussing<sup>96,99</sup> and if they are balanced.

## Comparisons Across Studies

Prevalence figures of inaccurate Web sites differ across different domains, eg, diet and nutrition sites (45.5%<sup>26</sup> and 88.9%<sup>64</sup> inaccurate information) vs cancer sites (4% for prostate cancer,<sup>46</sup> 5.1% for breast cancer,<sup>76</sup> 9% for English-language and 4% for Spanish-language breast cancer documents,<sup>14</sup> 6% for testicular cancer,<sup>46</sup> or 6.2% for Ewing Sarcoma<sup>16</sup>). While such prevalence figures may suggest that diet information on the Web is of poorer quality than cancer information, unadjusted comparisons across studies have to be made with

care, as at least 3 potential confounders should be considered.

First, results heavily depend on the rigor of the methodology used: studies that used personal opinion as a criterion standard found fewer inaccurate Web sites and more often came to a positive conclusion than studies using more rigorous criteria.

Second, many studies use the terms "Web site" and "Web page" interchangeably, making comparisons difficult.

Third, different and often poorly described sampling and selection strategies were used. While in theory a truly random sample of Web sites could be identified (by choosing random Internet protocol [IP] addresses<sup>100</sup>), this approach is not practical for identifying Web sites for a given topic. Thus, all studies used search engines, catalogs, or lists of popular Web sites. However, the choice of the search strategy may greatly confound the results. Most studies mimicked how consumers would search (although only 3 involved actual consumers), hand-selecting popular sites. As there is no consensus and little research on how a typical consumer searches,<sup>98</sup> studies used various strategies, mostly picking the top-ranked results from a search engine. As many search engines can rank the better sites first, the search tool could influence the results. Moreover, even the same search engine may give a different result if different sampling strategies (ie, which sites are picked) are used. For example, Suarez-Almazor et al<sup>83</sup> showed that, in the Webcrawler search engine, the first 20 ranked hits are more relevant, have less financial interests, contain less alternative therapies, and are more often nonprofit organizations, than sites ranked lower. Moreover, the selection of search terms may critically determine which Web sites are retrieved. For example, using the term "coronary heart disease," Eachus<sup>31</sup> found only 2 sites (among 110) provided by lay people. His conclusion that "the concern that the Internet would be a major source of low-quality health information, particularly that provided by unqualified members of the lay public, is not sup-

ported by [our] findings" may be confounded by the choice of search terms, as lay people might not necessarily use terms such as "coronary heart disease." Who devises the search strategy and conducts the search can also affect the quality of the retrieved sites. For example, in a comparison of results of a search devised by doctors with a search devised by information experts, Groot et al<sup>43</sup> observed differences in the credibility and accuracy scores of the retrieved samples.

### Comparison With Other Media

The quality of Web sites should be interpreted in the larger context of information in other media to determine whether the Web is "the beginning of an epidemic of misinformation or nothing more than a variation of what is endemic."<sup>2</sup> Many of the shortcomings detected likely are not specific to the Web and are also present in other media. For example, 2 of the 4 "erroneous information" elements found on 65 Web pages identified by Biermann et al<sup>16</sup> were found in the online version of the *Encyclopaedia Britannica*, which probably has the same inaccuracies in the printed edition. This issue, and the relatively low prevalence (6.2%) of inaccurate information, was generally ignored when this study was widely quoted in the lay media as evidence that the "Internet can be a quick link to bad health information."<sup>101</sup>

Studies assessing information in traditional media also frequently report high prevalences of inaccurate or incomplete information. In an early study, authors found 70% of health information broadcast on television to be inaccurate, misleading, or both.<sup>102</sup> In another study, authors rated as inaccurate 76% of the information about oral hygiene from television, 53% from magazines, and 12% from newspapers.<sup>103</sup> Another study of the popular press found 20% of the information on oral cancer to be a "mix of accurate and inaccurate information."<sup>104</sup> The proportion of inaccurate press reports on healthy eating was found to be 55% in free advertising newspapers, 28.9% in lifestyle magazines, 29.9% in general interest magazines,

17.5% in health magazines, and 14.1% in newspapers.<sup>105</sup> In another study, 50% of the advice in newspaper advice columns was rated inappropriate, with critical issues only partially covered or not covered at all in 76% of the articles, and 58% were rated unsafe or potentially dangerous.<sup>106</sup> Inquiries by telephone to libraries yielded a 3.6% rate of inaccurate information.<sup>107</sup>

The perceived quality problem on the Internet is not restricted to the health sector: a study investigating the quality of general scientific information found that 10% to 34% was inaccurate, 20% to 35% was misleading, and 48% to 90% was unreferenced.<sup>108</sup>

Very few studies in our sample directly compared Internet information with information found elsewhere. One study<sup>35</sup> evaluated both printed and Web-based patient education materials, concluding that "there was no significant difference between the ranges of scores [incorporating content, writing style, design, and readability] for Internet and non-Internet leaflets"; however, subjectively authors felt that "the overall quality of the Internet leaflets was more variable and the information less comprehensive." Two studies compared the readability of Internet information with the readability of printed information. One study<sup>33</sup> concluded that SMOG readability levels of Internet information were significantly higher compared with other (printed) patient information materials; the other study<sup>65</sup> observed no differences in the proportion of patient information written above the 9th grade level, but 87.5% of Web information materials vs only 14.3% of printed patient information, were deemed unsuitable based on the Suitability Assessment of Materials score (however, as the 8 Web-based patient education materials came from only 2 different Web sites, this result may not be representative). We identified only 1 study that compared the accuracy of advice obtained on the Internet (in a pharmacy newsgroup, hence this study was excluded from this review) with advice obtained in the real world (from drug information centers).<sup>109</sup>

In summary, the prevalence figures of inaccurate or incomplete Web information reported by the studies in this review are difficult to interpret or compare, are unlikely to be representative or generalizable, and must also be considered against the background of imperfect consumer health information in other media.

### Presentation Quality

The presentation criteria evaluated by most investigators are considered to be quality criteria because: (1) their presence is deemed ethical according to several codes of conduct for Web publishing<sup>110-114</sup> (eg, transparency and accountability criteria, such as disclosure of authorship); (2) they help to create context and to avoid misunderstandings (eg, disclosure of sponsorship, purpose); (3) they empower users to select the information that is best for them in their individual situation (eg, disclosure of target audience); (4) they empower users to validate the information themselves (eg, references, contact addresses); or (5) they may influence the accessibility of information or effectiveness of communication (eg, search capabilities, speed, design). In addition, several authors have attempted to establish whether technical criteria or other site characteristics may be used as predictors for content quality,<sup>11,42,47,58,87,91</sup> but the results have been conflicting and inconclusive. A number of studies found, perhaps unsurprisingly, that the length of a document is correlated with better content scores.<sup>42,58,87</sup> A few studies suggest that the source may be a predictor for content quality, with commercial sites often scoring lower than academic sites,<sup>59,74,75</sup> but the way in which content is evaluated may also influence this relationship. Other technical criteria influenced by the source include references, which are more often found on Web sites targeting medical professionals,<sup>23</sup> on academic<sup>46,79</sup> or educational<sup>89</sup> sites, or on sites owned by organizations.<sup>42</sup> Individual authors often are not disclosed on government sites<sup>14</sup> or sites of organizations and drug companies.<sup>42</sup> Given these complex relation-

ships, it seems unlikely that a simple scoring system could be developed that predicts content quality across domains and site populations.

As with content criteria, the ability to compare or pool results across studies is impeded by wide variations in sampling strategies, methods, and operational definitions even for the same quality criteria (eg, what constitutes "attribution"). Moreover, there is no consensus on the unit of evaluation (eg, should the update cycle be published on each Web page, or is a site-specific disclosure sufficient) or granularity with which information should be provided (eg, is it sufficient to post the year of last update, or should the exact date be given).

Despite this lack of consensus, some quality criteria are consistently given (eg, "ownership disclosure," with 99% of sites providing information) or not given (eg, "credentials of physicians not disclosed," with 97.5% not complying) (Table 2).

As with content criteria, presentation criteria should be put into context by comparison with non-Web-based information. While in our study 64% of Web sites failed to provide a date of update, investigators evaluating printed patient leaflets have found that 53%,<sup>35</sup> or, in another study, one-third,<sup>115</sup> did not include publication dates.

### Omissions

Since most authors used criteria derived from the print world,<sup>7</sup> we noted important omissions concerning Web-specific criteria. For example, very few studies evaluated the privacy policy or the possibility to encrypt confidential information, and few studies checked whether the target audience or the target country (which is important in a global medium<sup>116</sup>) were clearly disclosed. None of the studies tested usability<sup>117</sup> or accessibility (eg, compliance with guidelines of the Web Accessibility Initiative, ascertaining that the site is available to people with disabilities or with low-end technology), and only 1 study checked whether metadata were provided,<sup>76</sup> which could

greatly enhance the consumer's ability to select and filter information.<sup>6</sup>

### Measuring Progress of Health Communication on the Web

Among the public health objectives of the US Department of Health and Human Services is to "increase the proportion of health-related World Wide Web sites that disclose information that can be used to assess the quality of the site."<sup>118</sup> How progress in this area might be achieved or evaluated is an open question. A meta-analysis of cross-sectional studies is one possible way, but one must take into account the different methods used (or the methods must be standardized) before meaningful conclusions can be drawn. Longitudinal studies could be used to assess changes over time, using a consistent methodology either to identify and assess Web sites for a given topic (with the caveat that observed changes may be changes in the ability of users to find better sites, ie, improvements in search engine technology to rank better sites first) or to follow up a cohort of Web sites. Two studies have attempted a longitudinal approach, with conflicting results; one observed an improvement, and the other a deterioration, in quality.<sup>36,57,90</sup> The third possibility is to promote among site developers the use of machine-processable disclosure statements as Web site labels (meta-data),<sup>6,119</sup> which would allow automatic tracking and analysis of the proportion and characteristics of Web sites making such disclosure statements. This would also facilitate the development of intelligent systems able to guide users to trustworthy Web sites.<sup>120</sup>

### Conclusion

The epidemiology of consumer health information on the Web is an emerging research discipline at the intersection of medical informatics and public health. Many descriptive, cross-sectional studies have attempted to draw attention to perceived "outbreaks" of misinformation on the Web by estimating the proportion or prevalence of inadequate health informa-

tion. However, the individual's risk (R) of encountering an inadequate site on the Web is a function of both the proportion of inadequate information on the Web (P) and the inability (I) of the individual (or his tools) to filter the inadequate sites. Since studies usually report R, but not I, we cannot infer P, or adjust study results to make them comparable across domains or time. Even if we could know P, we would still not know how this measurement of "true" misinformation on the Web translates into health outcomes or critical incidents in a population. On an individual level, R can be reduced by improving the ability of the user to locate trustworthy sites or to filter the inadequate ones. Public e-health interventions such as MedCERTAIN<sup>121</sup> therefore strive to reduce P and I by increasing, for example, the proportion of health information providers making disclosure statements<sup>119</sup> and by empowering consumers to identify trusted sites through educational and technological innovations, including the possibilities of the semantic Web.<sup>120</sup>

Given the difficulty in interpreting descriptive studies without control or comparison groups, future studies should use analytic rather than descriptive approaches to investigate the relationship between quality markers and other variables, including outcomes, or to compare different Web site populations or media. Such studies are urgently needed to help in the ongoing process to develop methods and instruments to guide consumers to quality information and to identify factors that can be assessed to predict favorable patient outcomes.

Finally, studies evaluating content should also harness the potential of the Web as a source for qualitative data<sup>122</sup>: rather than getting bogged down by the question of *how much* information is inaccurate, one could analyze *where* and *why* gaps exist between evidence-based medicine and health information on the Internet, which may elicit a wealth of valuable data that may inform priorities for research, health communication, and education.



**Author Contributions:** Study concept and design: Eysenbach, Powell.

**Acquisition of data:** Eysenbach, Powell, Sa.

**Analysis and interpretation of data:** Eysenbach, Powell, Kuss.

**Drafting of the manuscript:** Eysenbach.

**Critical revision of the manuscript for important intellectual content:** Eysenbach, Powell, Kuss, Sa.

**Statistical expertise:** Eysenbach, Kuss.

**Obtained funding:** Eysenbach.

**Administrative, technical, or material support:** Sa.

**Study supervision:** Eysenbach.

**Related Information:** Online tables available at <http://jama.com>.

**Funding/Support:** Dr Eysenbach was partly funded by the European Union under the Action Plan for Safer Use of the Internet (<http://www.saferinternet.org>) Med-CERTAIN Project (<http://www.medcertain.org>). Mr Powell is funded by the NHS (South East) Research and Knowledge Management program.

## REFERENCES

- Eysenbach G. Consumer health informatics. *BMJ*. 2000;320:1713-1716.
- Coiera E. Information epidemics, economics, and immunity on the Internet. *BMJ*. 1998;317:1469-1470.
- Jadad AR, Gagliardi A. Rating health information on the Internet. *JAMA*. 1998;279:611-614.
- Kim P, Eng TR, Deering MJ, Maxfield A. Published criteria for evaluating health related web sites: review. *BMJ*. 1999;318:647-649.
- Wyatt JC. Commentary: measuring quality and impact of the World Wide Web. *BMJ*. 1997;314:1879-1881.
- Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet. *BMJ*. 1998;317:1496-1500.
- Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet. *JAMA*. 1997;277:1244-1245.
- Landis RJ, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
- Bohning D, Dammann UP. The application of methods of meta-analysis for heterogeneity modeling in quality control and assurance. In: Schulze R, Holling H, Bohning D, eds. *Meta-analysis: New Developments and Applications in Medical and Social Sciences*. Seattle, Wash: Hogrefe & Huber. In press.
- Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol*. 1991;44:1271-1278.
- Abbott VP. Web page quality: can we measure it and what do we find? *J Public Health Med*. 2000;22:191-197.
- Armstrong K, Schwartz JS, Asch DA. Direct sale of sildenafil (Viagra) to consumers over the Internet. *N Engl J Med*. 1999;341:1389-1392.
- Beredjikian PK, Bozontka DJ, Steinberg DR, Bernstein J. Evaluating the source and content of orthopaedic information on the Internet. *J Bone Joint Surg Am*. 2000;82:1540-1543.
- Berland GK, Elliott MN, Morales LS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. *JAMA*. 2001;285:2612-2621.
- Berland GK, Morales LS, Elliott MN, et al. Evaluation of English and Spanish Health Information on the Internet. Santa Monica: RAND/California HealthCare Foundation; 2001. Available at: <http://www.rand.org/publications/documents/interneteval/>. Accessibility verified April 25, 2002.
- Biermann JS, Golladay GJ, Greenfield ML, Baker LH. Evaluation of cancer information on the Internet. *Cancer*. 1999;86:381-390.
- Bloom BS, Iannacone RC. Internet availability of prescription pharmaceuticals to the public. *Ann Intern Med*. 1999;131:830-833.
- Bogenschutz MP. Drug information libraries on the Internet. *J Psychoactive Drugs*. 2000;32:249-258.
- Boyer EW, Shannon M, Hibberd PL. Web sites with misinformation about illicit drugs. *N Engl J Med*. 2001;345:469-471.
- Breul H, Boue L, Martin A. Panorama des ressources documentaires hospitalières françaises sur Internet [Overview of Internet information resources of interest to French hospitals]. *Semaine des Hopitaux*. 1999;75:257-262.
- Butzke I, Kramer KL. Orthopädie im World Wide Web: Universitäts-Homepage der G7-Staaten. *Z Orthop Ihre Grenzgeb*. 2000;138:413-418.
- Bykowski JL, Alora MB, Dover JS, Arndt KA. Accessibility and reliability of cutaneous laser surgery information on the World Wide Web. *J Am Acad Dermatol*. 2000;42:784-786.
- Chen LE, Minkes RK, Langer JC. Pediatric surgery on the Internet. *J Pediatr Surg*. 2000;35:1179-1182.
- Corpron CA, Lelli JL Jr. Evaluation of pediatric surgery information on the Internet. *J Pediatr Surg*. 2001;36:1187-1189.
- D'Alessandro DM, Kingsley P, Johnson-West J. The readability of pediatric patient education materials on the World Wide Web. *Arch Pediatr Adolesc Med*. 2001;155:807-812.
- Davison K. The quality of dietary information on the World Wide Web. *J Can Dietetic Assoc*. 1996;57:137-141.
- Davison K. The quality of dietary information on the World Wide Web. *Clin Perform Qual Health Care*. 1997;5:64-66.
- Diering CL, Palmer MH. Professional information about urinary incontinence on the World Wide Web. *J Wound Ostomy Continence Nurs*. 2001;28:55-62.
- Doupi P, van der Lei J. Rx medication information for the public and the WWW: quality issues. *Med Inform Internet Med*. 1999;24:171-179.
- Dracos A, Seta M. Egeszegugyi informaciok az Interneten laikusoknak: olaszorszagi tapasztalatok. [Health information for the lay person on the Internet: a survey of the Italian experience]. *Tudomanyos es Muszaki Tajekoztatás*. 1998;45:118-120.
- Eachus P. Health information on the Internet. *Int J Health Prom Educ*. 1999;37:30-33.
- Ellamushi H, Narenthiran G, Kitchen ND. Is current information available useful for patients and their families? *Ann R Coll Surg Engl*. 2001;83:292-294.
- Estrada CA, Hryniewicz MM, Higgs VB, Collins C, Byrd JC. Anticoagulant patient information material is written at high readability levels. *Stroke*. 2000;31:2966-2970.
- Eysenbach G. Online prescribing of sildenafil (Viagra) on the World Wide Web. *J Med Internet Res*. 1999;1:e10. Available at: <http://www.jmir.org/1999/2/index.htm>. Accessibility verified March 25, 2002.
- Fitzmaurice DA, Adams JL. A systematic review of patient information leaflets for hypertension. *J Hum Hypertens*. 2000;14:259-262.
- Frasca D, Malezieux R, Mertens P, Neidhardt JPH, Voiglio EJ. Review and evaluation of anatomy sites on the Internet (updated 1999). *Surg Radiol Anat*. 2000;22:107-110.
- Galimberti A, Jain S. Gynaecology on the Net. *J Obstet Gynaecol*. 2000;20:297-299.
- Gillies M. Access to the information should be made easier. *BMJ*. 2000;321:47-48.
- Gillois P, Colombet I, Dreau H, Degoulet P, Chatterlier G. A critical appraisal of the use of Internet for calculating cardiovascular risk. *Proc AMIA Symp*. 1999;775-779.
- Gordon JB, Barot LR, Fahey AL, Matthews MS. The Internet as a source of information on breast augmentation. *Plast Reconstr Surg*. 2001;107:171-176.
- Graber MA, Roller CM, Kaebler B. Readability levels of patient education material on the World Wide Web. *J Fam Pract*. 1999;48:58-61.
- Griffiths KM, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. *BMJ*. 2000;321:1511-1515.
- Groot D, Riet G, Khan KS, Misso K. Comparison of search strategies and quality of medical information of the Internet: a study relating to ankle sprain. *Injury*. 2001;32:473-476.
- Harmon D, Duggan M, Flynn N. Anaesthesia on the World Wide Web. *Anaesthesia*. 2000;55:728-729.
- Hatfield CL, May SK, Markoff JS. Quality of consumer drug information provided by four Web sites. *Am J Health Syst Pharm*. 1999;56:2308-2311.
- Hellawell GO, Turner KJ, Le Monnier KJ, Brewster SF. Urology and the Internet: an evaluation of Internet use by urology patients and of information available on urological topics. *BJU Int*. 2000;86:191-194.
- Hernández-Borges AA, Macías-Cervi P, Gaspar-Guardado MA, Torres-Álvarez de Arcaya ML, Ruiz-Rabaza A, Jiménez-Sosa A. Can examination of WWW usage statistics and other indirect quality indicators help to distinguish the relative quality of medical web-sites? *J Med Internet Res*. 1999;1:e1. Available at: <http://www.jmir.org/1999/1/e1/index.htm>. Accessibility verified March 25, 2002.
- Hersh WR, Gorman PN, Sacherek LS. Applicability and quality of information for answering clinical questions on the Web. *JAMA*. 1998;280:1307-1308.
- Hoffman-Goetz L, Clarke JN. Quality of breast cancer sites on the World Wide Web. *Can J Public Health*. 2000;91:281-284.
- Impicciatore P, Pandolfini C, Casella N, Bonati M. Possiamo fidarci delle informazioni sanitarie al pubblico che viaggiano su Internet? l'esempio della gestione domiciliare della febbre nel bambino. [Reliability of public-oriented health care resources in the World Wide Web: the case of the home management of fever in children]. *Ricerca e Pratica*. 1997;13:67-75.
- Impicciatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the World Wide Web. *BMJ*. 1997;314:1875-1879.
- Jiang YL. Quality evaluation of orthodontic information on the World Wide Web. *Am J Orthod Dentofacial Orthop*. 2000;118:4-9.
- Kihlstrom LC. Evaluating pharmacy benefit management information on the Internet. *Manag Care Interface*. 2001;14:64-68.
- Latthe PM, Latthe M, Khan KS. Quality of medical information about menorrhagia on the worldwide web. *BJOG*. 2000;107:39-43.
- Latthe M, Latthe PM, Charlton R. Quality of information on emergency contraception on the internet. *Br J Fam Plann*. 2000;26:39-43.
- Latthe PM, Latthe M, Khan KS. Quality of information on female sterilisation on the Internet. *J Obstet Gynaecol*. 2000;20:167-170.
- Li L, Irvin E, Guzman J, Bombardier A. Surfing for back pain patients. *Spine*. 2001;26:545-557.
- Libertiny G, Perkins JM, Magee TR, Galland RB. Varicose veins on the internet. *Eur J Vasc Endovasc Surg*. 2000;20:386-389.
- Lissman TL, Boehnlein JK. A critical review of internet information about depression. *Psychiatr Serv*. 2001;52:1046-1050.
- Mallory C. What's on the Internet? services for women affected by HIV and AIDS. *Health Care Women Int*. 1997;18:315-322.
- Martínez-López JF, Ruiz-Crespo EJ. Internet: calidad de la información sobre cirugía [Internet: quality of information available on orthopaedic surgery and traumatology]. *Revista de Ortopedia y Traumatología*. 1998;42:469-473.
- Maugans TA, McComb JG, Levy ML. The internet as a pediatric neurosurgery information resource. *Pediatr Neurosurg*. 1998;28:186-190.

63. McClung HJ, Murray RD, Heitlinger LA. The Internet as a source for current patient information. *Pediatrics*. 1998;101:e2.
64. Miles J, Petrie C, Steel M. Slimming on the Internet. *J R Soc Med*. 2000;93:254-257.
65. Murphy PW, Chesson AL, Berman SA, Arnold CL, Galloway G. Neurology patient education materials. *J Neurosci Nurs*. 2001;33:99-104.
66. Oermann MH, Wilson FL. Quality of care information for consumers on the Internet. *J Nurs Care Qual*. 2000;14:45-54.
67. O'Mahony B. Irish health care Web sites: a review. *Ir Med J*. 1999;92:334-337.
68. Ogushi Y, Tatsumi H. *Research on the Analysis of the Current State of the Provision and Use of Health Information Provided Through a New Technology Medium* [in Japanese]. Tokyo, Japan: Health & Welfare Ministry Research Group; 2000.
69. Tatsumi H, Mitani H, Haruki Y, Ogushi Y. Internet medical usage in Japan. *J Med Internet Res*. 2001;3:e12. Available at: <http://www.jmir.org/2001/1/e12>. Accessibility verified April 24, 2002.
70. Pandolfini C, Impicciatore P, Bonati M. Parents on the web. *Pediatrics*. 2000;105:e1.
71. Payne C, Miller K. The quality of footcare information on the Internet. *Australas J Podiatr Med*. 2000;34:63.
72. Peroutka SJ. Analysis of Internet sites for headache. *Cephalalgia*. 2001;21:20-24.
73. Roberts JR, Spooner SA. Pediatric Internet resources. *Arch Pediatr Adolesc Med*. 1997;151:592-597.
74. Sacchetti P, Zvara P, Plante MK. The Internet and patient education—resources and their reliability. *Urology*. 1999;53:1117-1120.
75. Sandvik H. Health information and interaction on the Internet. *BMJ*. 1999;319:29-32.
76. Shon J, Musen MA. The low availability of meta-data elements for evaluating the quality of medical information on the World Wide Web. *Proc AMIA Symp*. 1999;945-949.
77. Sing A, Salzman JR, Sing D. Problems and risks of unsolicited e-mails in patient-physician encounters in travel medicine settings. *J Trav Med*. 2001;8:109.
78. Smith M, Gertz E, Alvarez S, Lurie P. The content and accessibility of sex education information on the Internet. *Health Educ Behav*. 2000;27:684-694.
79. Soot LC, Moneta GL, Edwards JM. Vascular surgery and the Internet. *J Vasc Surg*. 1999;30:84-91.
80. Stausberg J, Fuchs J. Die chirurgische Fachabteilung im World Wide Web: Tribut an den Zeitgeist oder Informationsdrehscheibe? [Surgical specialty department in the World Wide Web: tribute to contemporary life style or information network?]. *Chirurg*. 2000;71:472-477.
81. Stausberg J, Fuchs J, Husing J, Hirche H. Health care providers on the World Wide Web. *Med Inform Internet Med*. 2001;26:17-24.
82. Stone TW, Jumper JM. Information about age-related macular degeneration on the Internet. *South Med J*. 2001;94:22-25.
83. Suarez-Almazor ME, Kendall CJ, Dorgan M. Surfing the Net—information on the World Wide Web for persons with arthritis. *J Rheumatol*. 2001;28:185-191.
84. Tamm EP, Raval BK, Huynh PT. Evaluation of the quality of self-education mammography material available for patients on the Internet. *Acad Radiol*. 2000;7:137-141.
85. Thompson JM, Howard JE. HMO Web sites. *Manag Care Q*. 2000;8:33-41.
86. Tu F, Zimmerman NP. It is not just a matter of ethics: a survey of the provision of health disclaimers, caveats, and other health-related alerts in consumer health information on eating disorders on the Internet. *Int Inf Libr Rev*. 2001;32:325-339.
87. Tülp JC, Gerds T, Neugebauer S. Myoarthropathien des Kausystems: Beurteilung der Qualität von Patienteninformationen im Weltweiten Netz. *Z Arztl Fortbild Qualitatssich*. 2001;95:539-547.
88. Neugebauer S, Tülp JC. Vorsicht Internet! Webseiten über myoarthropathien des kausystems [Attention Internet! Web sites on myoarthropathies of the mastication system]. *Schweiz Monatsschr Zahnmed*. 2001;111:298-311.
89. Veronin MA, Ramirez G. The validity of health claims on the World Wide Web. *Am J Health Promot*. 2000;15:21-28.
90. Voiglio E, Frasca D, Malezieux R, Moreau S, Rodier MN, Neidhardt JPH. Prospecting and evaluation of the anatomy sites on the internet. *Surg Radiol Anat*. 1999;21:65-68.
91. von Danwitz F, Baehring T, Scherbaum W. *Verbreitung, Gestaltung und Qualität von Deutschsprachigen World Wide Web Seiten zum Diabetes Mellitus*. Düsseldorf, Germany: Deutsches Diabetes Forschungsinstitut; 1999.
92. Willems M, Bouvy ML. Gezondheidsinformatie op Internet: ze adviseren gewoon de spullen die op voorraad zijn. *Pharmaceutisch Weekblad*. 2001;136:607-609.
93. Willems M, Bouvy ML. Gezondheidsinformatie op het internet : de kwaliteit van websites vergeleken. *Medisch Contact*. 2001;56:673-676.
94. Wilson FL, Baker LM, Brown-Syed C, Gollop C. An analysis of the readability and cultural sensitivity of information on the National Cancer Institute's Web site: CancerNet. *Oncol Nurs Forum*. 2000;27:1403-1409.
95. Wright B, Williams C, Partridge I. Management advice for children with chronic fatigue syndrome: a systematic study of information from the Internet. *Ir J Psychol Med*. 1999;16:67-71.
96. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health*. 1999;53:105-111.
97. Delamothe T. Quality of websites: kitemarking the west wind. *BMJ*. 2000;321:843-844.
98. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the World-Wide-Web? *BMJ*. 2002;324:573-577.
99. Dalton L, Gartenfeld E. Evaluating printed health information for consumers. *Bull Med Libr Assoc*. 1981;69:322-324.
100. Lawrence S, Giles CL. Accessibility of information on the web. *Nature*. 1999;400:107-109.
101. Internet can be quick link to bad health information. *CNN*. July 29, 1999. Available at: <http://www.cnn.com/HEALTH/9907/29/internet.health>. Accessibility verified April 25, 2002.
102. Smith FA, Trivax G, Zuehlke DA, Lowinger P, Nghiem TL. Health information during a week of television. *N Engl J Med*. 1972;286:516-520.
103. Frazier PJ, Jenny J, Ostman R, Frenick C. Quality of information in mass media: a barrier to the dental health education of the public. *J Public Health Dent*. 1974;34:244-257.
104. Canto MT, Kawaguchi Y, Horowitz AM. Coverage and quality of oral cancer information in the popular press: 1987-98. *J Public Health Dent*. 1998;58:241-247.
105. Institut für Ernährungswissenschaften der Universität Wien. *Österreichischer Ernährungsbericht*. Vienna, Austria: Bundesministerium für Arbeit, Gesundheit und Soziales und der Bundesministerin für Frauenangelegenheiten und Verbraucherschutz; 1998.
106. Molnar FJ, Man-Son-Hing M, Dalziel WB, et al. Assessing the quality of newspaper medical advice columns for elderly readers. *CMAJ*. 1999;161:393-395.
107. Paskoff BM. Accuracy of telephone reference service in health sciences libraries. *Bull Med Libr Assoc*. 1991;79:182-188.
108. Allen ES, Burke JM, Welch ME, Rieseberg LH. How reliable is science information on the web? *Nature*. 1999;402:722.
109. Seaboldt JA, Kuiper R. Comparison of information obtained from a Usenet newsgroup and from drug information centers. *Am J Health Syst Pharm*. 1997;54:1732-1735.
110. Risk A, Dzenowagis J. Review of Internet health information quality initiatives. *J Med Internet Res*. 2001;3:e28. Available at: <http://www.jmir.org/2001/4/e28>.
111. Boyer C, Selby M, Scherrer JR, Appel RD. The Health on the Net Code of Conduct for medical and health Websites. *Comput Biol Med*. 1998;28:603-610.
112. Hi-Ethics Group. *Health Internet Ethics: Ethical Principles For Offering Internet Health Services to Consumers*. 2000. Available at: <http://www.hiethics.org/Principles/>. Accessibility verified March 25, 2002.
113. e-Health Ethics Initiative. e-Health Code of Ethics. *J Med Internet Res*. 2000;2:e9. Available at: <http://www.jmir.org/2000/2/e9/>. Accessibility verified March 25, 2002.
114. Winker MA, Flanagan A, Chi-Lum B, et al. Guidelines for medical and health information sites on the Internet. *JAMA*. 2000;283:1600-1606.
115. Coulter A, Entwistle V, Gilbert D. Sharing decisions with patients: is the information good enough? *BMJ*. 1999;318:318-322.
116. Eysenbach G, Diepgen TL. Labeling and filtering of medical information on the Internet. *Methods Inf Med*. 1999;38:80-88.
117. Nielsen J. *Usability Engineering*. San Diego, Calif: Morgan Kaufmann; 1993.
118. US Department of Health and Human Services. *Healthy People 2010: Understanding and Improving Health*. 2nd ed. Washington, DC: Office of Disease Prevention and Health Promotion, US Dept of Health and Human Services; 2000. Publication 017-001-00547-9. Available at: <http://www.health.gov/healthypeople/document/>. Accessibility verified March 25, 2002.
119. Eysenbach G, Köhler C, Yihune G, Lampe K, Cross P, Brickley D. A metadata vocabulary for self- and third-party labeling of health web-sites. *Proc AMIA Annu Fall Symp*. 2001;169-173.
120. Eysenbach G. An ontology of quality initiatives and a model for decentralized, collaborative quality management on the (semantic) World-Wide-Web. *J Med Internet Res*. 2001;3:e34. Available at: <http://www.jmir.org/2001/4/e34>. Accessibility verified April 24, 2002.
121. Eysenbach G, Köhler C, Yihune G, Lampe K, Cross P, Brickley D. A framework for improving the quality of health information on the world-wide-web and bettering public (e-)health. *Medinfo2001, Proceedings of the Tenth World Congress on Medical Informatics, London, England, September 2001*. Amsterdam, Holland: IOS Press; 2001:1450-1454.
122. Eysenbach G, Till JE. Ethical issues in qualitative research on Internet communities. *BMJ*. 2001;323:1103-1105.

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\*

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Abbott, <sup>11</sup> 2000	NR (accepted October 1999)	40 Sites providing information on MMR vaccine	A (a posteriori, "reflects the most current research," based on MEDLINE searches, textbooks, and an expert consultation) C ("balance: stating side effects and contraindications along with advantages") T (16 items, partly referring to "page aesthetics") R (Flesch Reading Ease Score, calculated using WordPerfect v8) D (use of relevant graphics; overall aesthetics rating) S (author is "recognized authority" and "qualified to publish the document")	1	No	38	30	Negative
Armstrong et al, <sup>12</sup> 1999	April 14-May 7, 1999	77 Sites selling sildenafil without requiring a prescription or a visit to a physician	C (a priori, checked for example whether sites required or offered medical evaluation, what information was required from consumers' medical history [5 items], what information was provided about sildenafil [5 items]) T (eg, disclaimers and warnings present, qualification of physician provided)	1?	No	50	14	Negative
Beredjickian et al, <sup>13</sup> 2000	NR (published November 2000)	175 Sites providing information on carpal tunnel syndrome	"Soot score": A (therapy classified as "conventional, unconventional, or misleading," a posteriori, compared against "textbooks and literature") C (31 items with different weights, leading to an "information score" with 100 points maximum) T (authorship identifiable)	2	Yes	38	50	Negative
Berland et al, <sup>14,15</sup> 2001	October 18-30, 2000+ November 6-13, 2000	18 English and 7 Spanish sites covering breast cancer, childhood asthma, depression, obesity	A (100 a priori elements on 3-point scale ["mostly incorrect," "mostly correct," "completely correct"]) C (100 a priori elements rated on 3-point scale ["not covered," "minimally covered," "more than minimally covered"]) T (2 items: authors and affiliations and credentials disclosed, date material created or updated) R (Fry Readability Graph (FRG), SMOG grading formula (English only), Lexile Framework, Flesch-Kincaid)	3-4 (AC) per condition	Yes	62	95	Neutral

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Biermann et al, <sup>16</sup> 1999	NR (submitted December 1998)	371 (170 relevant) pages about Ewing Sarcoma, 65 (or 70) evaluated for accuracy	A (a posteriori, compared against textbook) T (reference source listed)	2	No	50	43	Negative
Bloom and Iannaccone, <sup>17</sup> 1999	February-March 1999	46 Web sites offering prescription drugs to consumers	T (eg, whether sites reveal their geographic location or specific address of consulting physicians)	1?	No	38	0	Negative
Bogenschutz, <sup>18</sup> 2000	NR (October 1998?)	3 Sites providing information on psychoactive drugs	AC (a posteriori, based on comparison with literature)	1	No	25	25	Positive
Boyer et al, <sup>19</sup> 2001	May 24, 2001	7 Partisan Web sites promulgating information about illicit drugs	A (a posteriori evaluation against personal opinion of experts whether sites make potentially harmful recommendations for the management of the adverse effects of illicit drugs)	>1	Yes	12	57	Negative
Breul et al, <sup>20</sup> 1999	1998	68 Sites of French health care facilities	T (20 a priori items)	1?	No	31	0	Neutral
Butzke and Kramer, <sup>21</sup> 2000	August 3, 1998–May 15, 1999	136 Sites of university departments of orthopedics, traumatology, and rheumatology from G7 countries	T (18 a priori items) D (part of "qualitative criteria," but no results reported)	1?	No	62	0	Negative
Bykowski et al, <sup>22</sup> 2000	August 1997	40 Sites containing information about cutaneous laser surgery	C (8 a priori items) T (disclaimer, references)	1?	No	62	14	Negative
Chen et al, <sup>23</sup> 2000	NR (published August 2000)	141 Sites providing information on 4 pediatric surgery topics	A (a posteriori, based on personal judgement of 2 surgeons, score system) C (a posteriori, based on personal judgement of 2 surgeons, score system) T (2 a priori items: "accountability criteria," references)	2	No	38	20	Negative
Corpron and Lelli, <sup>24</sup> 2001	NR (presented Oct 28, 2000)	8 Sites about intersex anomalies	A (a posteriori, based on 2 textbooks) T (references to source of information)	1?	No	31	14	Negative

(continued)



**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
D'Alessandro et al, <sup>25</sup> 2001	March 2000	100 (Final sample: 89) documents from pediatric patient education sites	R (Flesch Reading Ease score and Flesch-Kincaid reading level with MS Word 98; Fry formula and SMOG handcalculated)	1-2	Partially	50	50	Negative
Davison, <sup>26</sup> 1996 Davison, <sup>27</sup> 1997	February 1996	167 "Sites" (pages?) containing information about nutrition, food, and diet	A (11 a priori items, from guideline)	2	Yes	56	86	Negative
Diering and Palmer, <sup>28</sup> 2001	June–July 1999	15 Sites on urinary incontinence from professional organizations targeted for health care providers	T (10 a priori items, including author/organization credentials, currency, references) A (a posteriori, evaluated by 2 reviewers, unclear whether determined by consensus) R (Flesch-Kincaid Grade Level)	2	No	56	29	Neutral
Doupi and van der Lei, <sup>29</sup> 1999	February–June 1998	14 Sites offering comprehensive information on prescription medication; on each of these sites, 6 drug information articles evaluated	C (36 a priori items, based on guidelines) T (15 items, eg, disclosure of authorship, objectivity, disclosure, and currency)	1?	No	12	43	Negative
Dracos and Seta, <sup>30</sup> 1998	Summer 1997	30 Italian Web sites with health information for patients	T (4 Silberg criteria: authorship disclosed, references, sponsorship disclosed, dates) R (subjective rating of readability on a scale "very good," "good," "not sufficient")	1?	No	50	0	Negative
Eachus, <sup>31</sup> 1999	NR (published 1999)	86 Coronary heart disease sites	S (1-8 points depending on the "likely quality of the site")	1?	No	50	0	Positive
Ellamushi et al, <sup>32</sup> 2001	NR (published July 2001)	150 Sites about 5 different neurosurgery conditions and procedures	A ("usefulness: structured treatment of a disease or procedure aimed towards patients and their families, accurate")	1?	No	25	0	Positive
Estrada et al, <sup>33</sup> 2000	NR (submitted May 25, 2000)	9 Patient information documents about atrial fibrillation and warfarin from 6 different sites	R (SMOG, Flesch-Kincaid Grade Level)	N/A	Partially	0	50	Negative
Eysenbach, <sup>34</sup> 1999	March 1999	10 cyberpharmacies selling sildenafil	C (5 a posteriori items, checking contraindications on the prescribing form asked) R (consumer terms used)	1?	No	62	0	Negative

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Fitzmaurice and Adams, <sup>35</sup> 2000	NR (submitted July 1999)	42 Sites and 19 printed patient information leaflets on hypertension	AC ("marks were allocated for content," 8 a priori criteria, from consensus; listed for non-Internet leaflets only—unclear whether Web pages were evaluated using the same criteria) T (date provided, table of contents) R (Gunning Fog Index, subjective marks for writing style) D (section headings present, at least 12-point font)	2 (or 3, if disagreement >10% in score)	No	25	60	Neutral
Frasca et al, <sup>36</sup> 2000	June 20, 1999–September 15, 1999	48 Anatomy sites	CTD (same method as in Voiglio et al <sup>30</sup> : raters graded sites on a 0-5 scale in the categories navigability, illustration, presentation, and text; the latter being defined as "value and amount of text" without further criteria, giving a maximum total score of 20)	1?	No	50	0	Neutral
Galimberti and Jain, <sup>37</sup> 2000	May 20–June 1, 1999	26 Sites about hysterectomy	AC (5 broad a priori items from information leaflet produced by Royal College of Obstetricians and Gynaecologists) T (7 technical criteria based on HON principles) R (subjective rating whether "it was felt to be clearly presented in a legible way and in plain English")	1?	No	50	40	Neutral
Gillies, <sup>38</sup> 2000	NR (published July 2000)	292 Pages from 126 unique sites providing information on cancer	A (NR) C (NR) T (authority, scope, completeness, disclosure, accuracy, validity, objectivity, uniqueness, currency, audience, accessibility, navigation, functionality, links, interactivity) D ("aesthetic features")	1	No	25	10	Neutral
Gillois et al, <sup>39</sup> 1999	NR (published 1999)	8 Sites providing cardiovascular risk prediction tools	A (a posteriori, "information was based on some explicit evidence") T (42 a priori items) D (Visual Analog Scale)	3-4	Yes	38	57	Negative
Gordon et al, <sup>40</sup> 2001	November 21, 1999	41 Sites providing information about breast augmentation	AC (3 surgeons evaluated accuracy of descriptions of procedural details and whether they are "limited in quantity," as well as description of complications, against personal opinion) T (provision of interactivity)	3	No	50	15	Negative
Graber et al, <sup>41</sup> 1999	NR (submitted November 1998)	50 Sites relevant for patient education (32 topics)	R (Flesch reading score, Flesch-Kinkaid reading level)	N/A	N/A	25	N/A	Negative
Griffiths and Christenson, <sup>42</sup> 2000	March 1999	21 Sites providing information about depression	AC (43 a priori items from guideline, including 5 core items, plus 17 further "issues," also a global score on a 10-point scale) T (9 a priori criteria according to Silberg et al <sup>7</sup> )	2	Yes	50	90	Negative

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Groot et al, <sup>43</sup> 2001	NR (accepted February 8, 2001)	36 Sites with information about diagnosis and treatment of ankle sprain	A (10 a priori items, based on guideline and systematic reviews, each item scored 0-2 points) T/"Credibility score" (9 items: source disclosed, context, currency, utility, editorial review process, hierarchy of evidence, statement of original source, disclaimer, omissions noted) [no details provided on how currency or utility, for example, were evaluated]	1	No	50	57	Negative
Harmon et al, <sup>44</sup> 2000	NR (published July 2000)	120 Anesthesia sites	AC (34 a priori items in 5 categories, each scored as complete, inadequate, not mentioned, or incorrect)	1?	No	50	20	Negative
Hatfield et al, <sup>45</sup> 1999	October 1997–November 1997	4 Sites offering comprehensive information on prescription medication, on each of these sites 30 drug information articles evaluated	A (3 a priori items, from textbooks) C (22 a priori items) T (site sponsorship, authors and contributors, references, dates of most recent updates, ease of use)	1-3?	No	38	40	Neutral
Hellawell et al, <sup>46</sup> 2000	NR (accepted June 12, 2000)	50 Sites providing information about prostate cancer and 50 sites about testicular cancer	A (a posteriori, rated "unconventional" or "conventional" compared with textbooks and literature) C (Soot-score, 5 broad weighted items rated on a scale 0-10) T (references)	2	No	25	40	Positive
Hernández-Borges et al, <sup>47</sup> 1999	March–April 15, 1998	363 Sites related to pediatrics	T (time since last update, counter present, author provided) S (impact factor of Web site author)	1?	No	50	0	Negative
Hersh et al, <sup>48</sup> 1998	NR (published October 1998)	639 Pages retrieved in an attempt to answer 50 clinical questions	T (9 items, eg, disclosure of authorship, credentials, affiliation, funding, date posted, attribution)	1	No	19	25	Negative
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	November 1998–June 1999	136 Sites providing information about breast cancer	T (Silberg criteria, privacy disclaimer)	2	Partially	50	50	Negative
Impicciatore et al, <sup>50,51</sup> 1997	December 1996	41 Pages containing information about home management of children with fever	AC (5 a priori items, from guideline/textbook)	1?	No	75	60	Negative
Jiang, <sup>52</sup> 2000	March 20, 1998	70 Sites providing information about orthodontics	A (a posteriori, personal opinion on whether information correct, questionable or incorrect) T (Silberg criteria: authorship, attribution, currency, disclosure)	1?	No	62	0	Neutral

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Kihlstrom, <sup>53</sup> 2001	"Late 1999"	71 US pharmacy benefit management sites	T (18 technical criteria, including "information gathering with cookies," feedback mechanisms, "visitor-friendly print-size," currency disclosure, copyright notice, site map, search function, plug-in required, more than 2 levels beyond home page, site purpose evident, credentials of author, disclaimer for medical advice, confidentiality statement, advertising policy, separation of advertising from content, HON logo present)	1?	No	36	0	Neutral
Latthe et al, <sup>54</sup> 2000	December 22, 1998	9 Pages providing patient information on menorrhagia	AC (9 a priori items relating to the treatment, from guidelines) T (disclosure of author/source, date, editorial review process, indication of the strength of the evidence)	2	Yes	88	90	Negative
Latthe et al, <sup>55</sup> 2000	March 19, 1999	32 Sites (pages?) providing patient information on emergency contraception	AC (12, 5, and 5 a priori items relating to the 3 different methods of emergency contraception, from guidelines) T (disclosure of author/source, date, editorial review process, indication of the strength of the evidence)	2	Yes	75	90	Negative
Latthe et al, <sup>56</sup> 2000	June 24, 1999	12 Sites on female sterilization	AC (10 a priori items from "evidence-based clinical guidelines" of the Royal College of Obstetricians and Gynaecologists) T (disclosure of author/source, date, review process ["seal of approval"], indication of the strength of the evidence)	2	Yes	88	90	Negative
Li et al, <sup>57</sup> 2001	September 1996/January 1998/February 1999	Time series of 74/63/54 sites providing information on back pain	A ("evidence-based," a posteriori against guidelines, literature, personal opinion) AC (overall rating "poor/fair/good" concerning accuracy and comprehensiveness) T (references)	2/1/2	Yes	50	65	Negative
Libertiny et al, <sup>58</sup> 2000	September 10-21, 1999	41 Sites providing information on varicose vein surgery	"Soot score": C (6 broad items with different weights, leading to an "information score" with 100 maximum points) T (date, authors, references provided)	2	Yes	62	57	Negative
Lissman and Boehnline, <sup>59</sup> 2001	April 2000	176 Sites (including duplicates?) about depression treatment	AC (3 broad a priori items for therapy [sites had to mention consultation, medication, psychotherapy, without specifying which medications], and 14 a priori diagnostic criteria from <i>DSM-IV</i> leading to a score of 0-14 points with 1 point deducted for each erroneous statement)	1?	No	50	55	Negative

(continued)



**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Mallory, <sup>60</sup> 1997	November 1996	69 Sites related to HIV and AIDS	AC (a posteriori, rated on a 4-point scale from "poor" to "excellent," no explicit criteria)	1	No	50	10	Neutral
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	April 1, 1997	57 sites about rotator cuff rupture	A (2 a posteriori items assessing whether "the presented theory or technique is usually considered appropriate" and whether treatment limitations are discussed, against personal opinion) T (6 items, including authority of authors, updates, references, loading time) R (subjective assessment, whether style is easily readable) D (legibility: appropriate background, layout)	1?	No	62	0	Neutral
Maugans et al, <sup>62</sup> 1998	NR (paper received February 1998)	Unclear number of sites containing information about 10 different pediatric neurosurgery topics	A (a posteriori, compared against personal opinion)	1?	No	25	14	Positive
McClung et al, <sup>63</sup> 1998	NR (paper submitted June 1997)	70 Pages containing information on treatment of acute diarrhea in childhood (60 traditional + 10 alternative)	A (unspecified number of a priori items, from guideline)	1?	No	50	43	Negative
Miles et al, <sup>64</sup> 2000	NR (published May 2000)	45 Sites providing information on weight-loss diets	A (a posteriori, based on guidelines)	1?	No	38	36	Negative
Murphy et al, <sup>65</sup> 2001	NR (published April 2001)	43 Neurology patient education brochures, among them 8 from the WWW (2 different Web sites)	R (Fog Index) D (part of SAM) Suitability Assessment of Materials (SAM) score	1	No	0	25	Negative
Oermann and Wilson, <sup>66</sup> 2000	NR (published July 2000)	10 Sites containing quality-of-care information	R (Flesch-Kincaid of 10 documents, three 100-word paragraphs from each)	N/A	N/A	0	N/A	Neutral
O'Mahony, <sup>67</sup> 1999	October-December 1998	60 Sites from Ireland hosted by health care providers, educational or professional bodies	T (Silberg criteria: disclosure of authorship, affiliation and credentials, references and sources, disclosure of site ownership, sponsorship, etc; date posted, email posted) R (Flesch-Kincaid reading level) [done only for the 47 service information provides]	1?	No	31	0	Negative

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Ogushi and Tatsumi, <sup>68</sup> 2000 Tatsumi et al, <sup>69</sup> 2001	November 21 -December 5, 1999	1147 Sites of Japanese medical institutions (hospitals and clinics), 516 of which contained medical information that was evaluated for accuracy and design	A (a posteriori evaluation against personal opinion of specialists, rated as "no problem in the content," "minor problem," or "serious problem") T (statement of information providers' name, address, telephone, fax, privacy policy, unencrypted transmission for personal information, third-party seal) D (3-point scale ["good"—"difficult to see"])	1 or more	No	29	15	Negative
Pandolfini et al, <sup>70</sup> 2000	June 1997 -January 1998	19 Pages containing information about cough in children	A (6 a priori items, from guidelines) C (4 a priori items, from MEDLINE search) T (6-item checklist: lists authors, authors credentials, lists references, relevant links, statement of "not substitute for professional care," currency)	1?	No	62	50	Negative
Payne and Miller, <sup>71</sup> 2000	NR (published 2000)	75 Sites about podiatrics (flat feet, corns, heel spurs, bunions)	T (aggregate rating score 1-6 incorporating author and qualification disclosure, references, date provided, links available, information not biased)	1?	No	38	0	Negative
Peroutka, <sup>72</sup> 2001	January 2000	51 Headache pages	Aggregate score (maximum, 100), containing A (a posteriori, no clear criteria, maximum 20 points), C (a posteriori, 20 points maximum for "clinical content," "positively associated with amount of information") T (20 points maximum for references, 20 points maximum for author, date of creation or last edit, sponsor, disclaimer) D (maximum 20 points, incorporating ease of use, presence of an overview etc)	1?	No	71	0	Negative
Roberts and Spooner, <sup>73</sup> 1997	December 1996	300 Sites and mailing lists related to podiatry	S (Web sites rated for "reliability of the author" on a 5-point scale)	1	No	33	25	Neutral
Sacchetti et al, <sup>74</sup> 1999	January 1, 1998 -August 31, 98	61 Sites containing patient education on sildanefil	Self-developed composite score system, incorporating accuracy, comprehensiveness, and objectivity, each scored from 1-4. A (a posteriori, compared against literature/rater consensus, 4-point scale "statement incorrect," "some incorrect," "correct," "correct, referenced") C (4-point scale: "discusses isolated/several/most/all issues," no explicit items specified, a posteriori)	2	No	62	30	Negative

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Sandvik, <sup>75</sup> 1999	Spring 1998	75 Sites, cyberdoctors and newsgroups providing advice on female urinary incontinence	Self-developed composite score system, incorporating technical criteria and completeness: T (ownership, navigability, interactivity, balance, currency, source, authorship) AC (14 a priori broad topic items, each rated whether they are comprehensively explained, briefly explained, mentioned, or not mentioned. Unclear whether the comprehensiveness scale implies accuracy. Criterion standard unclear)	1	No	12	30	Positive
Shon and Musen, <sup>76</sup> 1999	NR (published 1999)	97 Sites containing information on breast cancer treatment	A (a posteriori, based on personal opinion) T (authorship, attribution, currency, and disclosure)	1?	No	44	0	Positive
Sing et al, <sup>77</sup> 2001	July 1999	73 Sites of airlines	T (e-mail addresses present, feedback tested)	1?	No	33	0	Negative
Smith et al, <sup>78</sup> 2000	November 1997	41 Sites providing information on sex education	C (10 items)	3 or 1	No	62	36	Neutral
Soot et al, <sup>79</sup> 1999	NR (submitted September 15, 1998)	146 Sites containing information on vascular surgery (aortic aneurysm, carotid surgery, leg ischemia)	A (therapy classified as "conventional, unconventional, or misleading," a posteriori, compared against personal opinion and FDA approval) C (6 broad topic items, no specific content requirements. Each item scored with up to 10 points and weighted with a item-specific factor from 1 to 3, leading to an "information score" with maximum 100 points) T (references given)	C: 2 A: 1	Yes	38	40	Negative
Stausberg and Fuchs, <sup>80</sup> 2000	September 1999	184 Presentations of German surgical departments	T (11 items, including "last update," "navigation," and 9 further items deemed to be important on home pages of surgical departments, eg, calendar of events, research publications, job offers) C (a posteriori, comprehensiveness of medical information rated "very good: diseases and therapeutic interventions explained in detail"—"sufficient: described shortly"—"insufficient") D (layout, on scale very good, sufficient, insufficient)	1-2	Yes	29	36	Negative

(continued)

**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Stausberg et al, <sup>81</sup> 2001	Unclear (either March 29, 2000, or May 29, 2000)	171 Presentations of German surgical departments	T (11 items, including "last update," "navigation," and 9 further items deemed to be important on home pages of surgical departments, eg, calendar of events, research publications, job offers) C (a posteriori, comprehensiveness of medical information rated "very good: diseases and therapeutic interventions explained in detail"—"sufficient: described shortly"—"insufficient") D (layout, on scale very good, sufficient, insufficient)	6	Yes	29	43	Negative
Stone and Jumper, <sup>82</sup> 2001	December 1998	80 Sites providing information on age-related macular degeneration	A (a posteriori, against literature, classified as "conventional, experimental, alternative")	1?	No	50	14	Negative
Suarez-Almazor et al, <sup>83</sup> 2001	May 1998	286 Web pages on 205 sites about rheumatoid arthritis	A (a posteriori, against personal opinion "alternative = not taught widely at US medical schools or generally available at US hospitals")	1-2	No	50	21	Negative
Tamm et al, <sup>84</sup> 2000	16 March 1998	38 Sites providing advice on screening mammography	A (1 a priori item, from guideline) T (Silberg criteria: disclosure of authorship, ownership, date of publication, or update provided, attribution)	3	No	62	71	Negative
Thompson and Howard, <sup>85</sup> 2000	March 2000	21 HMO Web sites (of the 25 largest HMOs)	C (5 a priori items, "shown in the literature to be most important to consumers")	1?	No	29	29	Neutral
Tu and Zimmerman, <sup>86</sup> 2001	NR (covers information published in 1998)	97 Sites providing information on eating disorders, anorexia, bulimia	T (8 items, mainly disclaimers and caveats)	1-3	Partially	38	50	Negative
Türp et al, <sup>87</sup> 2001 Neugebauer and Türp, <sup>88</sup> 2001	May 30-June 9, 2000	47 Web pages on myoarthropathies of the mastication system	A (a posteriori, 5-point score system) T (DISCERN) (partly C: 1 item of DISCERN refers to balance)	1?	No	50	14	Negative
Veronin and Ramirez, <sup>89</sup> 2000	December 1998–May 1999	184 Sites (98 claims) about the herbal remedy Opuntia	A (a posteriori comparison against primary literature, including extensive database searches and quality assessment of 51 scientific reports) T (references listed, disclaimer present)	1?	No	62	14	Negative
Voiglio et al, <sup>90</sup> 1999	February 2–June 10, 1998	52 Anatomy sites	ACTD (raters graded sites on a 0-5 scale in the categories navigability, illustration, presentation, and text; the latter being defined as "value and amount of text" without further criteria, giving a total score of maximum 20)	1 (-4?)	No	62	0	Neutral
von Danwitz et al, <sup>91</sup> 1999	October 1999	39 Sites about diabetes (German)	T (13 a priori "objective" and 12 a priori "subjective" items) D (visual design, as part of subjective evaluation)	5	Yes	50	75	Negative

(continued)



**Online Table A.** Included Studies Assessing Quality of Health Information on Web Sites or Web Pages\* (cont)

Study, Year	Period Conducted	Study Domain, Population	Evaluation Method and Criteria	No. of Raters	Interrater Reliability Determined?	S-score	E-score	Authors' Conclusion
Willems and Bouvy, <sup>92,93</sup> 2001	January-February 2001	23 Dutch health portals, including 6 cyberpharmacies	AC (a priori? accuracy and completeness were evaluated by 3 raters: "with a standard questionnaire the quality of 8 medical topics were tested against Dutch standards"—no further information given) T (technical criteria, including source of information clear, target audience clear, search functionality, all rated on 5-point scale ["very good"—"good"—"satisfactory"—"moderate"—"bad"]) R (subjective rating on "use of language") D ("clarity of design")	T:2 A:3 R,D:9-66	No	12	80	Negative
Wilson et al, <sup>94</sup> 2000	October 1997-July 1998	49 Documents from NCI's CancerNet Web site	R (Flesch-Kincaid reading level, measured with WinWord 95 v7.0); cultural sensitivity (5 yes/no questions)	2	Yes	20	75	Negative
Wright et al, <sup>95</sup> 1999	June 1997	13 Sites containing treatment information on chronic fatigue syndrome (CFS) in children	AC (accuracy rated a posteriori, compared against literature; completeness rating against 7 broad a priori criteria)	1?	No	50	20	Negative

\*S-score and E-score reflect the percentage of applicable study quality criteria (see Table 1) that were fulfilled by the respective study. NR indicates not reported; N/A, not applicable; A, accuracy; C, completeness; S, source; R, readability; T, technical; D, design (subjective); AC, accuracy and completeness evaluated together.

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\*

Study, Year	Description	Not Complying	Total	Not Complying, %	MA†
<b>Content of Site: Accuracy</b>					
Abbott, <sup>11</sup> 2000	Sites on MMR vaccine containing inaccurate or unbalanced information	24	40	60.0	
Beredjikian et al, <sup>13</sup> 2000	Sites on carpal tunnel syndrome classified as misleading (14%) or unconventional (9%)	41	175	23.0	
Berland et al, <sup>15</sup> 2001	English-language breast cancer sites rated not "completely correct" (average for several clinical elements)	N/A	10	9	
	English-language childhood asthma sites rated not "completely correct" (average for several clinical elements)	N/A	9	16	
	English-language depression sites rated not "completely correct" (average for several clinical elements)	N/A	10	25	
	English-language obesity sites rated not "completely correct" (average for several clinical elements)	N/A	10	14	
	Spanish-language breast cancer sites rated not "completely correct" (average for several clinical elements)	N/A	4	4	
	Spanish-language childhood asthma sites rated not "completely correct" (average for several clinical elements)	N/A	4	47	
	Spanish-language depression sites rated not "completely correct" (average for several clinical elements)	N/A	4	37	
	Spanish-language obesity sites rated not "completely correct" (average for several clinical elements)	N/A	4	32	
Biermann et al, <sup>16</sup> 1999	Web pages with clearly erroneous information about Ewing sarcoma	4	65‡	6.2	
Bogenschutz, <sup>18</sup> 2000	Inaccurate information on 3 drug sites concerning biological sources of psychoactive chemicals	0	13	0.0	
Boyer et al, <sup>19</sup> 2001	Partisan sites about illicit drugs making potentially harmful recommendations for the management of the adverse effects of illicit drugs	7	7	100	
Chen et al, <sup>23</sup> 2000	Pediatric surgery sites judged inaccurate (2 surgeons giving each 0-2 points for accuracy, composite score <3)	29	119	24	
Corpron and Lelli, <sup>24</sup> 2001	Sites about ambiguous genitalia offering misleading information or information not conforming to standard recommendations	3	8	37.5	
Davison, <sup>26</sup> 1996 Davison, <sup>27</sup> 1997	Noncompliance with Canadian nutritional guideline	76	167	45.5	
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations where "accuracy could not be determined"	3	15	20	
Galimberti and Jain, <sup>37</sup> 2000	Grossly incorrect or misleading statements made on sites about hysterectomy	3	26	11.5	
Gillois et al, <sup>39</sup> 1999	Cardiovascular risk prediction sites with "no valid use of information"	4	8	50.0	
Gordon et al, <sup>40</sup> 2001	Procedural details provided on breast augmentation sites inaccurate	3	28	11	
	Complications provided on breast augmentation sites inaccurate	2	19	11	
	Chat transcripts from 3 sites providing information on breast augmentation judged inaccurate	0	14	0	
Griffiths and Christensen, <sup>42</sup> 2000	Sites about depression contradicting or providing material inconsistent with 5 core items from guideline	12	21	58	
Hatfield et al, <sup>45</sup> 1999	Drug information sites with not 100% accurate drug information	1	4	25.0	
Hellawell et al, <sup>46</sup> 2000	Sites about prostate cancer providing "unconventional" information	2	50	4	
	Sites about testicular cancer providing "unconventional" information	3	50	6	
Jiang, <sup>52</sup> 2000	Sites about orthodontics rated incorrect	0	70	0	
Li et al, <sup>57</sup> 2001	Sites about back pain rated "not evidence-based" (September 1996)	45	73	61.6	
	Sites about back pain rated "not evidence-based" (February 1999)	40	54	74.1	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Content of Site: Accuracy (cont)</b>					
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Theories or techniques on pages about rotator cuff rupture rated inappropriate	11	57	19.3	
McClung et al, <sup>63</sup> 1998	Noncompliance with AAP guidelines about childhood diarrhea (A)	48	70	68.6	
Miles et al, <sup>64</sup> 2000	Sites about weight loss and diet giving unsound advice	40	45	88.9	
Ogushi and Tatsumi, <sup>68</sup> 2000	"Serious problem" on Japanese Web sites (additional 40 [7.8%] had a minor problem)	5	516	1	
Pandolfini et al, <sup>70</sup> 2000	Sites about cough in children giving more incorrect advice than correct advice	10	19	52.6	
Shon and Musen, <sup>76</sup> 1999	Pages about breast cancer judged misleading	5	97	5.1	
Soot et al, <sup>79</sup> 1999	Sites about vascular surgery classified as misleading (28%) or unconventional (4%)	16	50	32.0	
Stone and Jumper, <sup>82</sup> 2001	Sites about age-related macular degeneration providing "unconventional" information (14% alternative, 7% experimental)	17	80	21	
Suarez-Almazor et al, <sup>83</sup> 2001	Alternative therapy information ("not taught widely at US medical schools or generally available at US hospitals") on Web pages about rheumatoid arthritis	131	286	45.8	
Tamm et al, <sup>84</sup> 2000	Sites not reflecting US guidelines on mammography frequency	5	38	13.0	
Türp et al, <sup>87</sup> 2001	Pages about temporomandibular disorders where medical quality was rated "weak" (score MWQ<3)	40	47	85.1	
Veronin and Ramirez, <sup>89</sup> 2000	Claims about Opuntia found to have conflicting or contradictory reports about effects in the literature	3	33	9.1	
	Claims about Opuntia not supported by any reports in the literature	65	98	66.3	
Willems and Bouvy, <sup>93</sup> 2001	Accuracy on Dutch sites rated "bad" on 5-point scale	1	23	4.3	
Wright et al, <sup>95</sup> 1999	Sites on CFS in children giving unsupported etiologic explanations	4	13	30.8	
	Sites on CFS in children erroneously suggesting large amount of rest	11	13	15.4	
	Not reflecting treatment advice of a study about CFS in children	11	13	84.6	
<b>Content of Site: Completeness</b>					
Chen et al, <sup>23</sup> 2000	Pediatric surgery sites judged incomplete (2 surgeons giving each 0-2 points for completeness, composite score <3)	122	131	93	
Latthe et al, <sup>54</sup> 2000	Pages that did not mention all 9 information items deemed essential for treatment of menorrhagia	8	9	88.9	
Latthe et al, <sup>55</sup> 2000	Sites that did not mention all 22 information items deemed essential for emergency contraception	32	32	100.0	
Latthe et al, <sup>56</sup> 2000	Sites that did not mention all 10 information items deemed essential for female sterilization	11	12	91.6	
Lissman and Boehnlein, <sup>59</sup> 2001	Site about depression not containing information on all 3 treatment items regarded important by the investigator (consulting, medications, psychotherapy)	103	178§	57.9	
Pandolfini et al, <sup>70</sup> 2000	Sites which did not mention all 4 information items deemed essential about childhood cough	16	19	84.2	
Smith et al, <sup>78</sup> 2000	Sites about sex education not mentioning all 10 educational items	38	41	92.7	
Willems and Bouvy, <sup>93</sup> 2001	Completeness on Dutch sites rated "bad" (lowest rating on 5-point scale)	2	23	8.7	
<b>Content of Site: Standard of Care</b>					
Armstrong et al, <sup>12</sup> 1999	Sites selling sildanefil without offering or requiring an online medical evaluation	31	77	40	
Bloom and Iannacone, <sup>17</sup> 1999	Sites selling prescription drugs to consumers without requiring a prescription or online medical evaluation	9	46	19.6	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Content of Site: Completeness and Accuracy</b>					
Berland et al, <sup>15</sup> 2001	English-language breast cancer sites with minimal or no coverage or incorrect (average across elements)	N/A	10	47	
	English-language childhood asthma sites with minimal or no coverage or incorrect (range across elements)	N/A	9	64	
	English-language depression sites with minimal or no coverage or incorrect (range across elements)	N/A	10	56	
	English language obesity sites with minimal or no coverage or incorrect (range across elements)	N/A	10	63	
	Spanish language breast cancer sites with minimal or no coverage or incorrect (average across elements)	N/A	4	61	
	Spanish language childhood asthma sites with minimal or no coverage or incorrect (range across elements)	N/A	4	77	
	Spanish language depression sites with minimal or no coverage or incorrect (range across elements)	N/A	4	88	
	Spanish language obesity sites with minimal or no coverage or incorrect (range across elements)	N/A	4	85	
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy not containing accurately all 5 items considered important	16	26	61.5	
Impicciatore et al, <sup>51</sup> 1997	Noncompliance with guideline about childhood fever (A + C)	37	41	90.2	
Li et al, <sup>57</sup> 2001	Sites about back pain rated "poor" ("very limited and/or inaccurate information") overall (September 1996)	15	73	20.5	
	Sites about back pain rated "poor" ("very limited and/or inaccurate information") overall (February 1999)	20	54	37.0	
Voiglio et al, <sup>90</sup> 1999	Anatomy sites where text ("value and amount of text") was rated 0 (on a scale 0-5) (unclear whether accuracy was evaluated as well, no clear criteria for "value" and "amount" given)	5	52	9.6	
<b>Design and Aesthetics</b>					
Brul et al, <sup>20</sup> 1999	No logo on all pages of French health care facilities	15	68	22	
Stausberg and Fuchs, <sup>80</sup> 2000	"Extremely insufficient layout" on Web sites of German surgical departments	21	184	11.4	
Voiglio et al, <sup>90</sup> 1999	Anatomy sites where presentation ("visual aspect and legibility of the sites") was rated 0 (on a scale 0-5)	0	52	0	
Willems and Bouvy, <sup>93</sup> 2001	Clarity of design: sites rated "bad" on 5-point scale	0	23	0	
<b>Design and Aesthetics: Images and Illustrations</b>					
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Images on pages about rotator cuff rupture "not informative" ("abuse of images for decorative purposes")	13	57	22.8	
Voiglio et al, <sup>90</sup> 1999	Anatomy sites where illustration ("quality of the different pictures, diagrams and photographs contained in the site") was rated 0 (on a scale 0-5)	3	52	5.8	
<b>Disclosure: Authorship</b>					
Brul et al, <sup>20</sup> 1999	No identity of editor/author disclosed on French hospital sites	41	68	60	a
Doupi and van der Lei, <sup>29</sup> 1999	No name of an individual author	11	14	78.6	a
Dracos and Seta, <sup>30</sup> 1998	Author not provided on Italian sites	2	30	6.7	a
Griffiths and Christensen, <sup>42</sup> 2000	Authorship not disclosed on depression sites	8	21	38.1	a
Hatfield et al, <sup>45</sup> 1999	No authors given	1	4	25.0	a
Hernández-Borges et al, <sup>47</sup> 1999	No editor/author's name was given	226	363	62.3	a
Hersh et al, <sup>48</sup> 1998	No author indicated	435	629	69.2	a
Jiang, <sup>52</sup> 2000	Authorship not revealed on orthodontics sites	12	70	17.1	a
Latthe et al, <sup>54</sup> 2000	No name of the author or institutions name given on pages on menorrhagia	2	9	22.2	a
Latthe et al, <sup>55</sup> 2000	No clear mentioning of author on emergency contraception sites	4	32	12.5	a
Latthe et al, <sup>56</sup> 2000	No clear display of the author or institutions name given on sites on female sterilization	3	12	25	a

(continued)



**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Disclosure: Authorship (cont)</b>					
Libertiny et al, <sup>58</sup> 2000	Sites about varicose veins surgery where authors were unidentifiable	3	41	7.3	a
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Author can not be identified, including first and last name, on pages about rotator cuff rupture	11	57	19.3	a
O'Mahony, <sup>67</sup> 1999	No details of the authors of the content on Irish health information providers	12	18	66.7	a
Pandolfini et al, <sup>70</sup> 2000	No authors listed	5	19	26.3	a
Sandvik, <sup>75</sup> 1999	Authors not given	54	75	72	a
Shon and Musen, <sup>76</sup> 1999	No author name on page	78	97	80	a
Tamm et al, <sup>84</sup> 2000	Authors not given	28	38	73.7	a
von Danwitz et al, <sup>91</sup> 1999	No identity of editor/author disclosed	10	39	26	a
<b>Disclosure: Ownership</b>					
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy not disclosing the "nature of the organization behind it"	1	26	3.8	b
Griffiths and Christensen, <sup>42</sup> 2000	Ownership not disclosed on depression sites	1	21	4.8	b
Groot et al, <sup>43</sup> 2001	Source not disclosed	0	36	0	b
Sandvik, <sup>75</sup> 1999	Ownership not disclosed	0	75	0	b
Tamm et al, <sup>84</sup> 2000	Ownership not disclosed	0	38	0	b
<b>Disclosure: Location</b>					
Bloom and Iannacone, <sup>17</sup> 1999	Sites selling prescription drugs to consumers without revealing geographical location (city and country)	41	46	89.1	
<b>Disclosure: Person Responsible</b>					
Ogushi and Tatsumi, <sup>68</sup> 2000	Information providers name, including person who manages or is responsible for the Web site, not stated clearly on Japanese sites	298	1147	26.1	
<b>Disclosure: Source</b>					
Groot et al, <sup>43</sup> 2001	Original source not disclosed	11	36	30.6	c
Hatfield et al, <sup>45</sup> 1999	No source of drug information given	2	4	50.0	c
Türp et al, <sup>87</sup> 2001	DISCERN item 4: Not clear what sources of information were used to compile the publication (other than the author or producer) (Web pages on myoarthropathies of the mastication system)	36	47	76.6	c
Willems and Bouvy, <sup>93</sup> 2001	Origin of information clear: sites rated "bad" on 5-point scale	0	23	0	c
<b>Disclosure: Sponsorship/ Funding Source</b>					
Doupi and van der Lei, <sup>29</sup> 1999	No sponsorship disclosure on drug information sites	8	14	57.1	d
	No source of funding disclosed on drug information sites	6	14	42.9	d
Dracos and Seta, <sup>30</sup> 1998	Sponsorship and conflict of interest not disclosed on Italian sites	17	30	56.7	d
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy where source of funding, form of sponsorship, financial support, or publicity could not be clearly identified	16	26	61.5	d
Griffiths and Christensen, <sup>42</sup> 2000	Sponsorship not disclosed on depression sites	18	21	86	d
Hatfield et al, <sup>45</sup> 1999	No sponsorship disclosure on drug information sites	1	4	25.0	d
Hersh et al, <sup>48</sup> 1998	No financial and other support clearly indicated	623	629	99.0	d
<b>Disclosure: Advertising Distinct From Contents</b>					
Doupi and van der Lei, <sup>29</sup> 1999	No clear distinction between advertising and content on drug information sites	1	14	7.1	e
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites with advertising where "it was difficult to clearly differentiate the PBM site from the outside advertising"	8	8	100	e
Tu and Zimmerman, <sup>86</sup> 2001	Advertisements not distinct from content on eating disorder sites	27	97	27.8	e
<b>Disclosure: Partnership</b>					
Gillois et al, <sup>39</sup> 1999	Cardiovascular risk prediction sites with "no partnerships identified"	1	8	12.5	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Disclosure: Conflict of Interest</b>					
Hersh et al, <sup>48</sup> 1998	No disclosure of potential conflicts of interest	559	629	88.9	
<b>Disclosure: Statement of Purpose</b>					
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with "purpose not stated"	1	15	7	f
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites where "it was not possible to determine the sites purpose within two clicks of the mouse"	7	71	10	f
Tu and Zimmerman, <sup>86</sup> 2001	Sites on eating disorders not providing disclaimer on purpose	85	97	87.6	f
Türp et al, <sup>87</sup> 2001	DISCERN item 1: Aims not clear (Web pages on myoarthropathies of the mastication system)	26	47	55.3	f
<b>General Disclosures</b>					
Chen et al, <sup>23</sup> 2000	"Accountability criteria (disclosure of author's name, author's credentials, evidence for claims or copyright, Web site owner and the date the content was posted) not satisfied"	59	141	42	g
O'Mahony, <sup>67</sup> 1999	No disclosure of site ownership, sponsorship, advertising, commercial funding arrangements, or support on Irish health care Web sites.	49	60	81.7	g
Shon and Musen, <sup>76</sup> 1999	No disclosure on site	40	97	41	g
<b>Currency of Information: Date of Creation Disclosed</b>					
Breul et al, <sup>20</sup> 1999	No creation date on French hospital sites	55	68	81	h
Doupi and van der Lei, <sup>29</sup> 1999	No date on when information was written on drug information sites	14	14	100	h
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not disclose when page was created	117	136	86	h
Pandolfini et al, <sup>70</sup> 2000	Did not supply year of creation (range 93-97)	7	19	36.8	h
Türp et al, <sup>87</sup> 2001	DISCERN item 5: not clear when the information used or reported in the publication was produced (Web pages on myoarthropathies of the mastication system)	43	47	91.5	h
<b>Currency of Information: First Posted Disclosed</b>					
Doupi and van der Lei, <sup>29</sup> 1999	No date on when information was first posted on drug information sites	14	14	100	
von Danwitz et al, <sup>91</sup> 1999	"Date of publication" not clear	19	39	49	
<b>Currency of Information: Any of Creation/First Posted Disclosed</b>					
Latthe et al, <sup>54</sup> 2000	No date on creation or information posted given on pages on menorrhagia	4	9	44.4	
<b>Currency of Information: Last Update Disclosed</b>					
Breul et al, <sup>20</sup> 1999	No information on update on French hospital sites	44	68	64	i
Doupi and van der Lei, <sup>29</sup> 1999	No revision date on drug information sites	9	14	64.2	i
Dracos and Seta, <sup>30</sup> 1998	No dates disclosed on Italian sites	17	30	56.7	i
Hernández-Borges et al, <sup>47</sup> 1999	No information on last update given [of those which gave the information, they had been updated 47.5 weeks before (range, 0-395)]	189	363	52	i
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites which did not disclose when page was updated	91	136	66.9	i
Pandolfini et al, <sup>70</sup> 2000	Did not supply date modified	15	19	78.9	i
Stausberg et al, <sup>81</sup> 2001	Web presentations of German surgical departments with no indication of last update	106	171	62	i

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Currency of Information: Any of Creation/Update Disclosed</b>					
Hersh et al, <sup>48</sup> 1998	No date of posting/update given	517	629	82.9	j
Jiang, <sup>52</sup> 2000	Orthodontics sites without date posted or last updated	56	70	80	j
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites where "it was impossible to tell if the information was current"	61	71	85.9	j
Latthe et al, <sup>55</sup> 2000	No date of creation, posting, update or revision	15	32	46.9	j
Latthe et al, <sup>56</sup> 2000	No date of the original document or date when content was posted	7	12	58	j
Libertiny et al, <sup>58</sup> 2000	Sites about varicose veins surgery bearing no date (1 published in 1995, 2 in 1996, 2 in 1997, 2 in 1998, 9 in 1999)	25	41	61.0	j
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Date of publication not provided on pages about rotator cuff rupture	9	57	15.8	j
O'Mahony, <sup>67</sup> 1999	No details of the date content was posted or updated on Irish health care Web sites [of the 25 sites which gave this information, 10 sites (17%) had not been updated in the last 6 months]	35	60	58.3	j
Sandvik, <sup>75</sup> 1999	Date of publication or last update not provided	32	75	42.7	j
Shon and Musen, <sup>76</sup> 1999	Neither date nor copyright with date listed	61	97	36	j
Stausberg and Fuchs, <sup>80</sup> 2000	Web presentations of German surgical departments with no indication of creation or last update	43	184	23.3	j
Tamm et al, <sup>84</sup> 2000	Date of publication or last update not provided	9	38	23.7	j
<b>Currency of Information: Sites Modified in the Past 6 mo</b>					
Berland et al, <sup>15</sup> 2001	English dated Web sites not created or updated in past year	?	?	54	
	Spanish dated Web sites not created or updated in past year	?	?	83	
Griffiths and Christensen, <sup>42</sup> 2000	Depression sites not modified in the past month	12	21	57	
Stausberg et al, <sup>81</sup> 2001	Web presentations of German surgical departments with update more than 6 months ago (of those sites disclosing update date)	14	65	21.5	
Stausberg and Fuchs, <sup>80</sup> 2000	Web presentations of German surgical departments not updated in 1999 (as of September 1999) (of those sites disclosing update date)	31	141	22	
<b>Currency of Information: Technical Maintenance Date Disclosed</b>					
von Danwitz et al, <sup>91</sup> 1999	No date of technical maintenance disclosed	21	39	54	
<b>Currency of Information: Other</b>					
Butzke and Kramer, <sup>21</sup> 2000	Orthopedic department homepages "not updated sufficiently" (definition unclear)	?	42 (or 125?)	54	
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with "date not listed or information that was evaluated as out of date"	6	15	40	
Groot et al, <sup>43</sup> 2001	Sites not fulfilling the "currency" criterion [unclear how this was evaluated]	27	36	75	
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Information on pages about rotator cuff rupture is "static," ie, not periodically updated (unclear, how this was determined by raters)	30	57	52.6	
<b>Authority of Source: Credibility of Source (Source Rating)</b>					
Eachus, <sup>31</sup> 1999	"Less credible sources" among coronary heart disease sites	36	86	41.9	
Roberts and Spooner, <sup>73</sup> 1997	Reliability of author (no one responsible, suspicious)	45	300	15.0	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Authority of Source: Author Credentials</b>					
Butzke and Kramer, <sup>21</sup> 2000	No information on "authority and qualification of author" on orthopedic department home pages with medical information	31	42	73.8	k
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with "no author/organization credentials listed"	1	15	7	k
Doupi and van der Lei, <sup>29</sup> 1999	No/insufficient provision of authors' credentials	6	14	42.9	k
Griffiths and Christensen, <sup>42</sup> 2000	Authors' credentials not disclosed on depression sites	10	21	48.6	k
Hersh et al, <sup>48</sup> 1998	No credentials of authors indicated	509	629	80.9	k
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites where credentials of site owner could not be identified	93	136	68.4	k
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	No information on qualification or workplace of the author given on pages about rotator cuff rupture	10	57	17.5	k
Pandolfini et al, <sup>70</sup> 2000	No credentials of authors listed	7	19	36.8	k
Shon and Musen, <sup>76</sup> 1999	No credentials of authors listed	60	97	62	k
<b>Authority of Source: Credentials of Consulting Physicians</b>					
Armstrong et al, <sup>12</sup> 1999	Sites selling sildenafil offering online medical evaluation by a physician without providing specific information about the qualifications of the physician	27	27	100	l
Bloom and Iannacone, <sup>17</sup> 1999	Sites selling prescription drugs without revealing address, name, specialty, location, or qualification of consulting physician doing online medical evaluation	37	37	100	l
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites that offer medical advice without clearly identifying the credentials of the individuals offering such advice	15	17	88.2	l
<b>Authority of Source: Not Provided by Health Professional/Specialists</b>					
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy not provided by health professional, or no disclosure of this fact	10	26	38.5	
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Pages about rotator cuff tear where raters had an "objection" against the qualification of the author, ie, they are not physician specialists	10	57	17.5	
<b>Authority of Source: Authors' Affiliations</b>					
Doupi and van der Lei, <sup>29</sup> 1999	No/insufficient provision of authors' affiliations	4	14	28.6	m
Griffiths and Christensen, <sup>42</sup> 2000	Authors' affiliations not disclosed on depression sites	10	21	48.6	m
Hersh et al, <sup>48</sup> 1998	No affiliation of site clearly indicated	293	629	46.6	m
O'Mahony, <sup>67</sup> 1999	No authors' affiliations and credentials on Irish health information providers	14	18	77.8	m
Shon and Musen, <sup>76</sup> 1999	No affiliation on page	10	97	10	m
<b>Ease of Use: Navigation</b>					
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with "no easy navigation"	1	15	7	n
Sandvik, <sup>75</sup> 1999	Navigability insufficient (information scattered around)	1	75	1.3	n
Stausberg and Fuchs, <sup>80</sup> 2000	Insufficient navigation on Web sites of German surgical departments	31	184	16.8	n
Voiglio et al, <sup>90</sup> 1999	Anatomy sites where navigability ("aptitude of the site to allow good circulation within the site and toward other sites") was rated 0 (on a scale 0-5)	0	52	0	n

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Ease of Use: Navigation Depth</b> Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites which "were constructed using more than two levels beyond the homepage"	35	71	50	
<b>Ease of Use: Search Engine</b> Breul et al, <sup>20</sup> 1999	No search engine on French hospital sites	60	68	88	o
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with no internal search engine	7	15	47	o
Gillois et al, <sup>39</sup> 1999	Cardiovascular risk prediction sites with no search engine	5	8	62.5	o
<b>Ease of Use: Search Functionality</b> Gillois et al, 1999	Cardiovascular risk prediction sites with "no simple browsing"	1	8	12.5	
Willems and Bouvy, <sup>93</sup> 2001	Search functionality: sites rated "bad" on 5-point scale	1	23	4.3	
<b>User Support: Navigation Aids</b> Breul et al, <sup>20</sup> 1999	No navigation aids on French hospital sites	10	68	15	
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Pages about rotator cuff rupture with insufficient navigation aids (index, table of contents, or icons for navigation)	13	57	22.8	
<b>User Support: Site Map</b> Breul et al, <sup>20</sup> 1999	No site map on French hospital sites	10	68	15	
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites not offering a site map	4	71	5.6	
<b>User Support: Documentation</b> Breul et al, <sup>20</sup> 1999	No help files on French hospital sites	68	68	100	
<b>User Support: What's New</b> Breul et al, <sup>20</sup> 1999	No "What's New" section on French hospital sites	44	68	84	
<b>Accessibility and Availability: Speed</b> von Danwitz et al, <sup>91</sup> 1999	Speed: takes more than 10 sec to build up	5	39	95	
<b>Accessibility and Availability: Browser Compatibility</b> von Danwitz et al, <sup>91</sup> 1999	Sites not browser compatible	3	39	8	
<b>Accessibility and Availability: Language</b> Breul et al, <sup>20</sup> 1999	French hospital sites not available in more than 1 language	56	68	82	
<b>Accessibility and Availability: Cited in Search Engines</b> Breul et al, <sup>20</sup> 1999	French hospital sites not cited in a search engine	0	68	0	
<b>Accessibility and Availability: Cited on Other Sites</b> Breul et al, <sup>20</sup> 1999	French hospital sites not cited on other sites	0	68	0	
<b>Links: Links Present</b> Breul et al, <sup>20</sup> 1999	No "pertinent links" on French hospital sites	22	68	33	p
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with "no links or dead links"	0	15	0	p
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not provide links	29	136	21	p
Pandolfini et al, <sup>70</sup> 2000	Did not provide links relevant to cough	15	19	78.9	p
<b>Links: Broken Links</b> Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Pages about rotator cuff rupture with incorrect URLs (broken links)	11	57	19.3	

(continued)



**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Attribution and Documentation: References</b>					
Abbott, <sup>11</sup> 2000	No references on pages about MMR vaccine	34	40	85.0	q
Biermann et al, <sup>16</sup> 1999	No reference source (including no reference to peer-review) listed on Web pages about Ewing Sarcoma	57	165	34.5	q
Butzke and Kramer, <sup>21</sup> 2000	No information about the source of medical information on orthopedic department homepages	14	42	33.3	q
Bykowski et al, <sup>22</sup> 2000	Cutaneous laser surgery Web sites without references to peer-reviewed publications or research	40	40	100	q
Chen et al, <sup>23</sup> 2000	No "referral to a reliable source of information"	103	131	79	q
Corpron and Lelli, <sup>24</sup> 2001	Sites about ambiguous genitalia not "referencing their source of information"	8	8	100	q
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations with "references not cited"	3	15	20	q
Doupi and van der Lei, <sup>29</sup> 1999	No/insufficient references on drug information sites	12	14	85.7	q
Dracos and Seta, <sup>30</sup> 1998	No references on Italian sites	17	30	56.7	q
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy not "properly referenced"	6	26	23.1	q
Gillois et al, <sup>39</sup> 1999	Cardiovascular risk prediction sites with no references	3	7	42.9	q
Griffiths and Christensen, <sup>42</sup> 2000	References not given on depression sites	12	21	57.1	q
Hatfield et al, <sup>45</sup> 1999	No references on drug information sites	1	4	25.0	q
Hellawell et al, <sup>46</sup> 2000	Sites about prostate cancer not providing conventional references	37	50	74	q
	Sites about testicular cancer not providing conventional references	34	50	68	q
Hersh et al, <sup>48</sup> 1998	No sources, bibliography listed	552	629	87.8	q
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not "identify sources for information" (as opposed to those identifying experts with credentials)	95	136	69.9	q
Jiang, <sup>52</sup> 2000	Orthodontics sites with no references related to the content	60	70	85.7	q
Li et al, <sup>57</sup> 2001	Sites about back pain rated "not providing an independent way to verify info (eg, a reference list)" (September 1996)	46	73	63.0	q
	Sites about back pain rated "not providing an independent way to verify info (eg, a reference list)" (February 1999)	34	54	63.0	q
Libertiny et al, <sup>58</sup> 2000	Sites about varicose veins surgery with no "conventional references"	38	41	92.7	q
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Pages about rotator cuff rupture not providing a bibliography	32	57	56.1	q
O'Mahony, <sup>67</sup> 1999	No references or sources given on sites of Irish health information providers	13	18	72.2	q
Pandolfini et al, <sup>70</sup> 2000	No references	16	19	84.2	q
Sandvik, 1999	No indication of source	50	75	66.7	q
Shon and Musen, <sup>76</sup> 1999	References not listed (if references were applicable, which was the case in 20%)	66	97	68	q
Soot et al, <sup>79</sup> 1999	No conventional references on pages which are able to reference	19	108	17.6	q
	(No conventional references, in relation to all pages)	(57)	(146)	(39.0)	q
Tamm et al, 2000	No references	11	38	28.9	q
Veronin and Ramirez, <sup>89</sup> 2000	No references or single reference (n = 7) on sites providing information about Opuntia	152	184	82.6	q
von Danwitz et al, <sup>91</sup> 1999	No sources cited	7	39	18	q

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Attribution and Documentation: Further Information or Links</b>					
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy not providing "contact addresses, links, or sources for further information"	6	26	23.1	
<b>Attribution and Documentation: Balanced Evidence</b>					
Gordon et al, <sup>40</sup> 2001	Breast augmentation sites with photographs showing either only good (74%) or only bad results (11%) (none of the physician sites showed bad results)	16	19	84	
	Breast augmentation sites judged as being biased toward a particular technique	34	41	82	r
Sandvik, <sup>75</sup> 1999	Sites providing unbalanced information	15	75	20	r
Türp et al, <sup>87</sup> 2001	DISCERN item 6: Unbalanced or biased (Web pages on myoarthropathies of the mastication system)	29	47	61.7	r
<b>Attribution and Documentation: Omissions Noted</b>					
Groot et al, <sup>43</sup> 2001	Omissions not noted on sites about ankle sprain	31	36	86.1	
Tu and Zimmerman, <sup>86</sup> 2001	Omissions not noted on eating disorder sites	95	97	97.9	
<b>Attribution and Documentation: Limitations Noted</b>					
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Pages about new techniques or theories about rotator cuff rupture where limitations were not discussed	12	57	21.1	
<b>Intended Audience: Target Audience Disclosed</b>					
O'Mahony, <sup>67</sup> 1999	Irish health care Web sites that did not specify their target audience	46	60	76.7	
Willems and Bouvy, <sup>93</sup> 2001	Target audience clear: sites rated "bad" on 5-point scale	0	23	0	
<b>Intended Audience: Multiple Target Audiences</b>					
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites "targeting just one audience"	12	71	16.9	
<b>Readability: Readability Formulas</b>					
Abbott, <sup>11</sup> 2000	Flesch-Reading Ease Score, calculated using WordPerfect v8 (no results reported)				
Berland et al, <sup>15</sup> 2001	English sites: Mean Fry Readability Graph (FRG) reading grade level 13.2 (SD, 2.1) Mean SMOG 13.6 (SD, 0.9) Lexile Framework 11.7 (SD, 1.0)				
	Spanish sites: Mean FRG reading grade level 9.9 (SD, 2.5) Lexile Framework 10.0 (SD, 2.6)				
D'Alessandro et al, <sup>25</sup> 2001	89 Documents: Mean Flesch Reading Ease score was 57.0 Mean Fry Formula was 12.0 (12 <sup>th</sup> grade, 0 months of schooling) Mean SMOG was 12.2 Mean Flesch-Kincaid grade level was 7.1				
Diering and Palmer, <sup>28</sup> 2001	12 Documents from 11 sites on urinary incontinence from professional organizations: Flesch Kinkaid Grade Level ranged from 6.2 to 14.5 years (median 10.5)				
Estrada et al, <sup>33</sup> 2000	9 patient information documents about atrial fibrillation and warfarin from 6 different sites: Flesch Kinkaid Grade Level ranged from 8.6 to 11.8 (median 9.6)				
	SMOG Grade Level ranged from 11.0 to 15.0 (median 12.0) Internet brochures had a significantly higher score than printed brochures.				
Fitzmaurice and Adams, <sup>35</sup> 2000	Gunning Fog Index, subjective marks for writing style (no results reported).				
Graber et al, <sup>41</sup> 1999	50 patient education Web sites had a mean Flesch reading score of 47.1 (median 44.1; range 25.7-70.3) and a mean Flesch-Kinkaid score of 9.9 (median 10.2 = 10 <sup>th</sup> grade 2 <sup>nd</sup> month; range 6.1-12). Authors cite references that suggest that 70%-80% of patients could not comprehend information at such reading level				

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Readability: Readability Formulas (cont)</b>					
Murphy et al, <sup>65</sup> 2001	7/8 (87.5%) of Web information materials and 31/35 (88.6%) of printed patient information was written above 9 <sup>th</sup> -grade level (Fog index). The 8 patient education materials came from only 2 different sites, so that it is questionable whether this result is representative.	7	8	87.5	
Oermann and Wilson, <sup>66</sup> 2000	10 documents about quality-of-care information for consumers: Flesch-Kincaid Mean Grade Level Scores ranged from 7 <sup>th</sup> -12 <sup>th</sup> grade (median 8) " <i>The overall mean reading demands of four of the documents were higher than the recommended 8<sup>th</sup> level accepted for the public</i> " Most used active voice, minimal technical jargon, font size from 12-14 points. None used illustrations or culturally specific information.	4	10	40	
O'Mahony, <sup>67</sup> 1999	46 Irish information service provider had a mean Flesch-Kincaid score of 15.7 years, (SD was 2 and the median was 17). Seventy-six percent had a reading age in excess of 12 to 14 years.	35	46	76.1	
Wilson et al, <sup>94</sup> 2000	Mean Flesch-Kincaid reading level of 49 documents from PDQ was 12 <sup>th</sup> grade (SD, 2.91)				
<b>Readability: Subjective Ratings of Writing Style</b>					
Dracos and Seta, <sup>30</sup> 1998	Readability of Italian sites rated "not sufficient" (subjective rating on 3-point scale)	4	30	13	s
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy rated not "to be clearly presented in a legible way and in plain English with avoidance of technical terminology or jargon"	3	26	11.3	s
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	The writing style of pages about rotator cuff rupture is not "easy to read" (subjective rating yes/no)	10	57	17.5	s
Willems and Bouvy, <sup>93</sup> 2001	"Use of language": 5 Dutch sites rated "satisfactory" (worst rating given) on 5-point scale (vs 18 "good")	5	23	21.7	s
<b>Readability: Legibility (Technical)</b>					
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites where "the print size was too small to read easily"	44	71	62	
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	Inadequate background (no sufficient contrast) on pages about rotator cuff rupture	24	57	42.1	
	Document illegible due to "abuse of different fonts and character sizes" or bad formatting of the text on pages about rotator cuff rupture	9	57	15.8	
<b>Readability: Reading Levels Noted</b>					
D'Alessandro et al, <sup>25</sup> 2001	Documents from pediatric patient education sites where reading level was not noted on the document	89	89	100	
<b>Contact Addresses or Feedback Mechanism: Feedback Mechanisms Provided</b>					
Hatfield et al, <sup>45</sup> 1999	No contact possibility for author on drug information sites	4	4	100	t
Sandvik, <sup>75</sup> 1999	No possibility for interactivity (e-mail or feedback form)	0	75	0	t
von Danwitz et al, <sup>91</sup> 1999	No feedback possibility to Web master	9	39	23	t
	No contact possibility editor/author	9	39	23	t
<b>Contact Addresses or Feedback Mechanism: Postal Address</b>					
Breul et al, <sup>20</sup> 1999	No mailing address on French hospital sites	41	68	60	
Ogushi and Tatsumi, <sup>68</sup> 2000	Address of the information provider not stated clearly on Japanese sites of medical institutions (hospitals, clinics)	126	1147	11.2	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Contact Addresses or Feedback Mechanism: Telephone</b>					
Breul et al, <sup>20</sup> 1999	No telephone number on French hospital sites	46	68	67	
Ogushi and Tatsumi, <sup>68</sup> 2000	Telephone number not stated clearly on Japanese sites of medical institutions (hospitals, clinics)	145	1147	12.6	
<b>Contact Addresses or Feedback Mechanism: Fax</b>					
Breul et al, <sup>20</sup> 1999	No fax number on French hospital sites	51	68	75	u
Diering and Palmer, <sup>28</sup> 2001	Sites on urinary incontinence from professional organizations "not interactive" (meaning no e-mail address provided?)	3	15	20	u
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites not offering feedback and e-mail	9	71	12.7	u
Ogushi and Tatsumi, <sup>68</sup> 2000	Fax number of the information provider not stated clearly on Japanese sites of medical institutions (hospitals, clinics)	439	1147	38.3	u
Thompson and Howard, <sup>85</sup> 2000	HMO sites not containing "specific information and guidance on a how to contact a customer service representative"	0	21	0	u
<b>Contact Addresses or Feedback Mechanism: E-mail</b>					
Breul et al, <sup>20</sup> 1999	No e-mail address on French hospital sites	25	68	37	v
Butzke and Kramer, <sup>21</sup> 2000	No e-mail addresses present on orthopedic department homepages	11	136	8.1	v
Doupi and van der Lei, <sup>29</sup> 1999	No provision of contact address or e-mail on drug information sites	3	14	21.4	v
Gillois et al, <sup>39</sup> 1999	Cardiovascular risk prediction sites with no e-mail provided	6	8	75.0	v
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not provide e-mail (but telephone number)	17	136	12.5	v
O'Mahony, <sup>67</sup> 1999	Webmaster e-mail not clearly displayed throughout the site	31	60	51.7	v
Ogushi and Tatsumi, <sup>68</sup> 2000	E-mail address of the information provider not stated clearly on Japanese sites of medical institutions (hospitals, clinics)	276	1147	24.1	v
Sing et al, <sup>77</sup> 2001	No feedback e-mail address on airlines sites	18	73	24.7	v
<b>Contact Addresses or Feedback Mechanism: Feedback Tested</b>					
Butzke and Kramer, <sup>21</sup> 2000	No response to inquiry sent to e-mail address published on orthopedic department homepages	80	125	64.0	
Sing et al, <sup>77</sup> 2001	No response to unsolicited e-mails asking airlines for malaria prophylaxis recommendations	30	55	54.5	
<b>Miscellaneous: Disclaimers (General)</b>					
Bykowski et al, <sup>22</sup> 2000	No disclaimer on cutaneous laser surgery Web sites	37	40	92.5	w
Doupi and van der Lei, <sup>29</sup> 1999	No disclaimers on drug information sites	0	14	0	w
Groot et al, <sup>43</sup> 2001	No disclaimer	14	36	38.8	w
Pandolfini et al, <sup>70</sup> 2000	No disclaimer ("not a substitute for professional care")	15	19	78.9	w
Tu and Zimmerman, <sup>86</sup> 2001	No caveats on eating disorder sites	84	97	86.6	w
Veronin and Ramirez, <sup>89</sup> 2000	No disclaimer (eg, for "educational purposes only" or "not a substitute for medical advice") on sites about Opuntia	146	184	79.3	w
<b>Miscellaneous: Disclaimers (Specific)</b>					
Kihlstrom, <sup>53</sup> 2001	No disclaimer on pharmacy benefit management sites that offer medical advice saying that "medical advice provided is not designed or intended to replace the relationship between the visitor and the health care provider"	4	17	23.5	
Tu and Zimmerman, <sup>86</sup> 2001	No disclaimer of limitations on eating disorder sites	90	97	92.8	
	No disclaimer of scope on eating disorder sites	90	97	92.8	
	No disclaimer of authority on eating disorder sites	84	97	86.6	
	No disclaimer of currency on eating disorder sites	97	97	100	
Veronin and Ramirez, <sup>89</sup> 2000	No FDA-disclaimer ("this product has not been evaluated by the FDA") on sites about Opuntia	172	184	93.5	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Miscellaneous: Copyright Notes</b>					
Doupi and van der Lei, <sup>29</sup> 1999	No copyright notes on drug information sites	0	14	0	x
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not provide copyright note	55	136	40.4	x
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites not including a copyright notice	24	71	33.8	x
Shon and Musen, <sup>76</sup> 1999	No copyright notice	51	97	53	x
<b>Editorial Review Process/Editorial Board</b>					
Breul et al, <sup>20</sup> 1999	No editorial committee on French hospital sites	61	68	90	y
Griffiths and Christensen, <sup>42</sup> 2000	No editorial board on depression sites	16	21	76	y
Groot et al, <sup>43</sup> 2001	No editorial review process	27	36	75	y
Latthe et al, <sup>54</sup> 2000	Did not provide editorial review process	9	9	100	y
Latthe et al, <sup>55</sup> 2000	Did not provide editorial review process	32	32	100	y
<b>Miscellaneous: External Review Process</b>					
Latthe et al, <sup>56</sup> 2000	No 'seal of approval' from a credible individual or group as evidence of the review process	12	12	100	
Ogushi and Tatsumi, <sup>68</sup> 2000	No seal showing the trustworthiness of the site on Japanese sites of medical institutions (hospitals, clinics)	1142	1147	99.6	
<b>Miscellaneous: Evidence Hierarchy</b>					
Groot et al, <sup>43</sup> 2001	Did not show hierarchy of evidence	30	36	83.3	z
Latthe et al, <sup>54</sup> 2000	Did not show hierarchy of evidence	5	9	55.6	z
Latthe et al, <sup>55</sup> 2000	Did not show hierarchy of evidence	26	32	81.3	z
Latthe et al, <sup>56</sup> 2000	Did not show hierarchy of evidence	7	12	58	z
<b>Miscellaneous: Level of Evidence</b>					
Martinez-Lopez and Ruiz-Crespo, <sup>61</sup> 1998	"Simply expresses personal opinion" (as opposed to "relevant information") on pages about rotator cuff rupture	10	57	17.5	
<b>Miscellaneous: Imprint</b>					
Shon and Musen, <sup>76</sup> 1999	No editorial statement on site	66	97	68	
von Danwitz et al, <sup>91</sup> 1999	No imprint (editorial statement)	8	39	21	
<b>Miscellaneous: Site Statistics</b>					
Breul et al, <sup>20</sup> 1999	French hospital sites not providing user statistics	58	68	85	
Hernández-Borges et al, <sup>47</sup> 1999	Pediatric Web sites not providing a visit counter	354	363	90	
<b>Miscellaneous: Disclosure of Charges</b>					
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not provide information on financial charges associated with products or services associated with the site	122	136	89.7	
Thompson and Howard, <sup>85</sup> 2000	HMO sites not providing information on monthly cost to consumers on their Web site	17	21	81.0	
<b>Miscellaneous: Confidentiality/Privacy</b>					
Galimberti and Jain, <sup>37</sup> 2000	Sites about hysterectomy where "confidentiality of data relating to individual patients was not respected" (eg, showing patients' photographs and full names)	7	26	26.9	
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites not providing a "clear confidentiality statement"	53	71	74.6	
Ogushi and Tatsumi, <sup>68</sup> 2000	Japanese sites of medical institutions (hospitals, clinics) not having a privacy policy	1145	1147	99.8	
<b>Miscellaneous: Encryption/Security</b>					
Hoffman-Goetz and Clarke, <sup>49</sup> 2000	Breast cancer sites that did not "indicate that information sent by the user would be a secure transmission or had a disclaimer indicating that messages would not be secured"	129	136	94.9	
Ogushi and Tatsumi, <sup>68</sup> 2000	Japanese sites of medical institutions (hospitals, clinics) not using encryption for transmitting personal information	1147	1147	100	
<b>Miscellaneous: Cookies</b>					
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites sending a "cookie"	23	71	32	
<b>Miscellaneous: Metadata</b>					
Shon and Musen, <sup>76</sup> 1999	No metadata in source code	45	97	46	

(continued)



**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
<b>Miscellaneous: Other</b>					
Berland et al, <sup>15</sup> 2001	Percentage of Spanish-language Web pages with neither author nor date	?	?	44	
	Percentage of English-language Web pages with neither author nor date	?	?	9	
Groot et al, <sup>43</sup> 2001	"No context provided"	1	36	2.8	
Kihlstrom, <sup>53</sup> 2001	Pharmacy benefit management sites not carrying a HON logo	69	71	97.2	
	Pharmacy benefit management sites requiring additional plug-ins	9	71	13	
<b>Global Ratings</b>					
Ellamushi et al, <sup>32</sup> 2001	Sites about pallidotomy rated "not useful" (from first 30 hits found by a search engine)	7	30	23	
	Sites about lumbar discectomy rated "not useful"	10	30	33.5	
	Sites about hydrocephalus rated "not useful"	3	30	10.5	
	Sites about glioma rated "not useful"	11	30	36.5	
	Sites about carotid artery aneurysm rated "not useful"	7	30	23.5	
Groot et al, <sup>43</sup> 2001	"Utility" criterion not fulfilled [unclear how this was evaluated]	10	36	3.6	
Mallory, <sup>60</sup> 1997	HIV/AIDS sites rated "poor" (on a 4-point Likert scale: poor/fair/good/excellent)	2	69	3	
Peroutka, <sup>72</sup> 2001	Headache page rated "less than optimal", ie, receiving <51 points in the evaluation	41	51	80	
<b>Suitability Assessment of Materials (SAM)</b>					
Murphy et al, <sup>65</sup> 2001	Based on the Suitability Assessment of Materials (SAM) score, 7/8 (87.5%) of Web information materials, but only 5/35 (14.3%) of printed patient information was deemed unsuitable. The 8 patient education materials came from only 2 different sites, so that it is questionable whether this result is representative.	7	8	87.5	
<b>DISCERN¶</b>					
Türp et al, <sup>87</sup> 2001	DISCERN item 2: Aims not reached (Web pages on myoarthropathies of the mastication system)	24	47	51.1	
	DISCERN item 3: Not relevant	23	47	48.9	
	DISCERN item 7: Does not provide details of additional sources of support and information	41	47	87.2	
	DISCERN item 8: Does not refer to areas of uncertainty	38	47	80.9	
	DISCERN item 9: Does not describe how each treatment works	40	47	85.1	
	DISCERN item 10: Does not describe the benefits of each treatment	29	47	61.7	
	DISCERN item 11: Does not describe the risks of each treatment	45	47	95.7	
	DISCERN item 12: Does not describe what would happen if no treatment is used	31	47	66.0	

(continued)

**Online Table B.** Overview of Quality Criteria Used by Studies and Their Evaluation Results\* (cont)

Study, Year	Description	Not complying	Total	Not Complying, %	MA†
	DISCERN item 13: Does not describe how the treatment choices affect overall quality of life	34	47	72.3	
	DISCERN item 14: Is not clear that there may be more than one possible treatment choice	30	47	63.8	
	DISCERN item 15: Does not provide support for shared decision-making	42	47	89.4	
	DISCERN item 16: Overall DISCERN score $\leq 2$ (serious or extensive shortcomings)	27	47	57.4	

\*Presented are the absolute and relative figures of Web pages/Web sites not complying with the respective quality criterion as reported in the study. Quality criteria have been classified according to the categories used by Kim et al.<sup>4</sup> MMR indicates measles, mumps, rubella; N/A, not applicable; AAP, American Academy of Pediatrics; CFS, chronic fatigue syndrome; PBM, pharmacy benefit management; URL, uniform resource locator; HMO, health maintenance organization; FDA, US Food and Drug Administration; HIV, human immunodeficiency virus; and AIDS, acquired immunodeficiency syndrome.

†Letters indicate quality criteria that have been applied by more than 3 studies; thus, their results have entered the meta-analysis (Table 3).

‡Inconsistent data reported (reported as 65 or 70 in article).

§Authors did not exclude duplicates and may have counted many sites 2 or more times.

||Contains unrelated, dead, and/or duplicate sites.

¶Score  $\leq 2$  indicates "not complying."