Papers

Examination of instruments used to rate quality of health information on the internet: chronicle of a voyage with an unclear destination

Anna Gagliardi, Alejandro R Jadad

Abstract

Objective This study updates work published in 1998, which found that of 47 rating instruments appearing on websites offering health information, 14 described how they were developed, five provided instructions for use, and none reported the interobserver reliability and construct validity of the measurements.
Design All rating instrument sites noted in the original study were visited to ascertain whether they were still operating. New rating instruments were identified by duplicating and enhancing the comprehensive search of the internet and the medical and information science literature used in the previous study. Eligible instruments were evaluated as in the original study.
Results 98 instruments used to assess the quality of websites in the past five years were identified. Many of the rating instruments identified in the original study were no longer available. Of 51 newly identified rating instruments, only five provided some information by which they could be evaluated. As with the six sites identified in the original study that remained available, none of these five instruments seemed to have been validated.
Conclusions Many incompletely developed rating instruments continue to appear on websites providing health information, even when the organisations that gave rise to those instruments no longer exist. Many researchers, organisations, and website developers are exploring alternative ways of helping people to find and use high quality information available on the internet. Whether they are needed or sustainable and whether they make a difference remain to be shown.

Introduction

The quality of health information on the internet became a subject of interest to healthcare professionals, information specialists, and consumers of health care in the mid-1990s. Along with the rapid growth of healthcare websites came a number of initiatives, both academic and commercial, that generated criteria by which to ensure, judge, or denote the quality of websites offering health information. Some of these rating instruments took the form of logos resembling "awards" or "seals of approval" and appeared prominently on the websites on which they were bestowed.

In 1997 we undertook a review of "award-like" internet rating instruments in an effort to assess their reliability and validity.[1] We hypothesised that if the rating instruments were flawed they might influence healthcare providers or consumers relying on them as indicators of accurate information. Instruments were eligible for review if they had been used at least once to categorise a website offering health information and revealed the rating criteria by which they did so. The rating instruments were evaluated according to, firstly, a system for judging the rigour of the development of tools to assess the quality of randomised controlled trials[2] and, secondly, whether their criteria included three indicators suggested as appropriate for judging the quality of website content.[3 4] These indicators were authorship (information about authors and their contributions, affiliations, and relevant credentials), attribution (listing of references or sources of content), and disclosure (a description of website ownership, sponsorship, underwriting, commercial funding arrangements, or potential conflicts of interest). These criteria were selected for use in the original study because they could be rated objectively.

Our original study found that of 47 rating instruments identified, 14 described how they were developed, five provided instructions for use, and none reported the interobserver reliability and construct validity of the measurements. The review showed that many incompletely developed instruments were being used to evaluate or draw attention to health information on the internet.

The purpose of this study is to update the previous review of award-like rating instruments for the evaluation of websites providing health information and to describe any changes that may have taken place in the development of websites offering health information to practitioners and consumers with respect to the quality of their content.

Methods

We visited the websites describing each of the rating instruments noted in the original study to ascertain whether they were still operating. If internet service was disrupted for technical reasons or if sites were not

Graduate Department of Health Policy, Management and Evaluation, Faculty of Medicine, University of Toronto, Toronto, ON, Canada
Anna Gagliardi
senior research associate

Departments of Health Policy, Management and Evaluation, and Anaesthesia, University Health Network, University of Toronto, Toronto, ON, Canada
Alejandro R Jadad
professor

Correspondence to: A R Jadad, Director, Centre for Global eHealth Innovation, University Health Network, Toronto General Hospital, Fraser Elliott Building, 4th Floor, 190 Elizabeth Street, Toronto, ON M5G 2C4, Canada ajadad@uhnres. utoronto.ca

Table 1 Summary of criteria for rating instruments

Rating system	Health specific scope	Silberg et al[3] and Wyatt[4]			Moher et al[2]					Criteria changed from original study
		Authorship	Attribution	Disclosure	Type of instrument	No of items	Scale development	Reliability	Instructions	
Previously reviewed sites										
American Medical Association (ama-assn.org/ama/pub/category1/3952.html)	Y	√	√	√	U	8	NR	NR	NR	N
Argus Clearinghouse seal of approval (clearinghouse.net/ratings.html)	Y	NC	NC	NC	U	U	NR	NR	NR	N
GrowthHouse excellence award (growthhouse.org/award.html)	Y	NC	U	U	S (stars)	U	NR	NR	U	N
Health on the Net Foundation code of conduct (www.hon.ch/HONcode/Conduct.html)	Y	√	√	√	Logo	8	U	NR	Y	N
Medaille d'Or for website excellence (arachnid.co.uk/award/select.html)	N	NC	NC	NC	S (medals)	U	NR	NR	U	N
OncoLink's editors' choice awards (oncolink.upenn.edu/ed_choice/)	Y	NR	NR	NR	Logo	NR	NR	NR	NR	Y
Newly identified sites										
World wide web health awards (healthawards.com/wwwha/s2001Webawards/assessment.htm)	Y	NC	NC	NC	Logo	U	NR	NR	NR	—
HardinMD clean bill of health (lib.uiowa.edu/hardin/md/cbh.html)	Y	NC	NC	NC	Logo	U	NR	NR	NR	—
Nutrition Navigator among the best (navigator.tufts.edu)	Y	NC	√	NC	S	5	NR	NR	U	—
Pacific Bell knowledge network blue web'n (kn.pacbell.com/wired/blueWebn/rubric.html)	N	√	√	U	S	U	NR	NR	NR	—
(kn.pacbell.com/wired/blueWebn/rubric.html)										
Health Improvement Institute Aesculapius award for rating sites (hii.org)	Y	√	√	√	Logo	U	NR	NR	NR	—

Y=yes; N=no; S=scale; U=unclear; NR=not reported; NC=not considered; √=considered.

available on first visit, we attempted a connection on one further occasion.

The search strategies, inclusion and exclusion criteria, and techniques for data extraction were similar to those used in the original review.[1] We used the following sources to identify new rating instruments:

● A search to 7 September 2001 of Medline, CINAHL, and HealthSTAR (from December 1997) using [(top or rat: or rank: or best) and (internet or web) and (quality or reliab: or valid:)]

● A search of the databases Information Science Abstracts, Library and Information Science Abstracts (1995 to September 2001), and Library Literature (1996 to September 2001) using [(rat: or rank: or top or best) and (internet or web or site) and (health:)]

● A search to September 2001 using the search engines Lycos (lycos.com), Excite (excite.com), Yahoo (yahoo.com), HotBot (hotbot.com), Infoseek (go.com), Looksmart (looksmart.com), and Google (google.com) with [(rate or rank or top or best) and (health)]. Open Text (opentext.com) and Magellan (magellan.com), which were used in the first study, no longer function as internet search engines

● A review of messages about rating instruments and the quality of health related websites posted to the Medical Library Association listserv medlib-l (listserv.acsu.buffalo.edu/archives/medlib-l.html) and the Canadian Health Libraries Association listserv canmedlib-l (lists.mun.ca/archives/canmedlib.html)

● A search of the American Medical Informatics Association's 1998, 1999, 2000, and 2001 annual symposium programmes (www.amia.org) for mention of health information on the internet

● A search of the *Journal of Medical Internet Research* (September 1999 to September 2001) for mention of

evaluations of the quality of health information on the internet (www.jmir.org)

● A search of the online archive of the magazine *Internet World* (www.internetworld.com) (January 2000 to September 2001) for mention of health information on the internet.

We also reviewed relevant articles referenced in identified studies and links available on identified websites. We did not search the discussion list Public Communication of Science and Technology, which was consulted in the original study.

We stopped searching for rating instruments on 22 September 2001. Rating instruments were eligible for inclusion in the review if it was possible to link from their award-like symbol to an available website describing the criteria used by an individual or organisation to judge the quality of websites on which the award was bestowed. We excluded rating instruments from review if they were used only to rate sites offering non-health information or did not provide any description of their rating criteria. In contrast to the initial study, we did not contact the developers of rating instruments to request information about their criteria if it was not publicly available on their website.

We identified the website, group, or organisation that developed each eligible rating instrument, along with its web address. The two authors independently evaluated each rating instrument according to its validity (number of items in the instrument, availability of rating instructions, information on the development of rating criteria, and evaluation of interobserver reliability) and incorporation of the proposed criteria for evaluation of internet sites: authorship, attribution, and disclosure.[2-4]

Results

Fourteen rating instruments identified in the original study provided a description of their rating criteria and were therefore eligible for review. Six of these continued to function. Of the remaining eight instruments, four were no longer in operation and four had converted to a directory format. Table 1 summarises the review of the six functioning instruments. Our evaluation of one of these instruments, OncoLink's editors' choice awards, differed from that in the original study because the organisation does not provide information about the instrument on its website.

Of the 33 rating instruments identified in the original study that were not eligible for review, three continued to function. These were Best Medical Resources on the Web (priory.com/other.htm), Dr Webster's website of the day (drWebster.com), and HealthSeek quality site award (healthseek.com). None of these rating instruments revealed its rating criteria, and they therefore remained ineligible for review. Of the remaining rating instrument websites, 10 were no longer in operation, five had been subsumed by or merged with another organisation and had a different name or purpose, and 15 still offered a website but did not function as a rating instrument.

We newly identified 51 rating instruments. Eleven of these were identified as award-like symbols on a website offering health information, but the website of the organisation from which they originated was no longer operating (table 2). Of the remaining 40 rating instruments, 35 were associated with an active website but did not reveal the criteria by which they judge websites and were ineligible for evaluation (table 3). Five award sites discussed their evaluation criteria and were assessed (table 1). Although three of these five rating instruments exhibited one or more of the characteristics of authorship, attribution, and disclosure, none reported on the reliability and validity of the measurements or provided instructions on how to obtain the ratings.

Discussion

During the past five years, we have identified a total of 98 different rating instruments that have been used to assess the quality of websites. Many of the rating instruments identified in the original study were no longer available. Fifty one additional rating instruments have been developed since 1997, and many of these had also stopped functioning. Of 51 newly identified rating instruments, only five provided some information by which they could be evaluated. As with the six rating instrument sites identified in the original study that remained available, none of these seems to have been validated. Many incompletely developed rating instruments continue to appear on websites providing health information, even when the organisations that gave rise to them no longer exist. Surprisingly, many of these rating instruments, of questionable utility and without association to an operable entity, are featured on the US Department of Health and Human Services Healthfinder website (www.healthfinder.gov/aboutus/awards.htm), which uses a detailed and rigorous selection process for the development of its own content.

Table 2 Newly identified award sites not available

Rating instrument	Address
Computer Currents interactive link of the week award	currents.net
E-Medic Online medical award for excellence	emediconline.com
Eye on the Web selected site award	eyeontheWeb.com
Family Education Resource Network top family site	familytrack.com
Internet Voyager 5-star site	internetvoyager.com
Lesbianmoms and Gaydads site award winner	lesbianmoms.org
Nicecom nicelinks	nicecom.com
Smart Computing top website	smartcomputing.com
Starting Point choice award	www.stpt.com
USA Today hot site	www.usatoday.com
WebNet web rating	www.Webratings.net

Table 3 Newly identified available award sites not eligible for review

Rating instrument	Address
100hot	100hot.com
Achoo site of the week	achoo.com
Aids Awareness recognition award	www.geocities.com/WestHollywood/3390/AARAWARD.HTM
Awesome Library editor's choice	awesomelibrary.org
Beagle WebPick	biomednet.com
BioMed Link	links.bmn.com
Brill's Content best of the web	inside.com
Complete idiot's guide to online health and fitness	fitnesslink.com
Fitness Partner Connection champion websites	primusWeb.com/fitnesspartner/library/features/tour0198.htm
Forbes best of the web	forbes.com/bow/
FSPronet site of the week	fspronet.com
Go Network website award	go.com
Golden Web Awards	goldenWebawards.com
GoTo.com editor's choice award	goto.com
GovSpot spotlight award	govspot.com
Hammer award	surgeongeneral.gov/todo/pressreleases/HammerRel2.htm
Health Launchbase	health.launchbase.net
HealthLinks selects site	healthlinks.net
HotSheet featured site	hotsheet.com
Library Spot site of the month	libraryspot.com/refsiteofmonth0499.htm
Links2Go key resource award	links2go.com/award/Hospice
Mac's Picks recommended websites	2x2.co.nz/picks/
MedExplorer top rated category listings	medexplorer.com/toprated.dbm
Popular Science 50 best of the web	popsci.com
Rainbow award	www.gayamerica.com/awards/
RE Library pure gold award	relibrary.com
Suite 101 top 5 website	suite101.com
the1000.com webmaster select site	the1000.com
thegoodWebguide.co.uk recommended site	thegoodWebguide.co.uk
Third Age 701 special sites	thirdage.com
Top 100 health sites	www.health-top100.com/
Top 100 network	100.com/Top/Health
Web100	Web100.com/listings/health.html
World hottest 100 health websites	worldhot.com
Yahoo! Internet Life's 100 best sites for 2001	zdnet.com/techlife/

Our initial questions remain unanswered. Is it desirable or necessary to assess the quality of health information on the internet? If so, is it an achievable goal given that quality is a construct for which we have no gold standard? Some effort has been made to identify whether the presence of rating instrument awards influences consumers of health information,[5] but whether validated rating instruments would have an impact on the competence, performance, behaviour, and health outcomes of those who use them remains unclear.

Our search of the literature and the internet revealed that a large number of researchers, organisa-

Table 4 Initiatives to organise and identify valid health information on the internet

Organisation	Product	Description	Cost associated with use
US Department of Health and Human Services (www.healthfinder.com)	Healthfinder	Directory of health resources selected according to explicit criteria	—
Government of Australia (www.healthinsite.gov.au)	HealthInsite	Directory of health resources selected according to explicit criteria	—
National Health Service (nhsdirect.nhs.uk)	NHS Direct Online	Directory of health resources selected according to DISCERN criteria	—
Health Summit Working Group (hitiWeb.mitretek.org/hswg/)	Information quality tool	21 criteria by which consumers can evaluate websites	—
Health on the Net Foundation (www.hon.ch)	HON code of conduct	8 criteria to guide development of website content	—
Internet Healthcare Coalition (www.ihealthcoalition.org)	e-Health code of ethics	14 criteria by which consumers can evaluate websites	—
DISCERN on the Internet (discern.org.uk)	Questionnaire and user manual	16 criteria by which consumers can evaluate websites	—
Hi-Ethics Principles (www.hiethics.com)	E-Health seal	14 criteria to guide development of website content	$20 000 annual membership fee
American Accreditation HealthCare Commission (www.urac.org)	Health website accreditation programme	53 criteria to guide development of website content	Sliding scale based on revenue ranges from $3799 to $12 249
TRUSTe (www.truste.org)	Online seal ("trustmark") and mechanism for resolution of disputes	For consumers purchasing on line or providing personal information	Sliding scale based on revenue and number of brands ranges from $399 to $25 000 annually
Council of Better Business Bureaus (bbbonline.org)	Reliability seal and privacy seal programme plus mechanism for resolution of disputes	Online reliability standards to guide truthful advertising	Membership of Better Business Bureau; fees not disclosed.

tions, and website developers are exploring alternative ways to help people find and use high quality information available on the internet. Many reviews of healthcare information on the internet have been conducted, overall and for specific diseases or conditions.[6–12] Examination of over 90 reviews concluded that the validity of health information available on websites is highly variable across different diseases and populations, and is in many cases potentially misleading or harmful (G Eysenbach, personal communication, 2001). Several organisations, including government and non-profit entities, have developed criteria by which to organise and identify valid health information (table 4). Other groups, such as the OMNI Advisory Group for Evaluation Criteria (omni.ac.uk) and the Collaboration for Critical Appraisal of Information on the Net (www.medcertain.org), are refining technical mechanisms by which users of the internet can easily locate quality health information in a transparent manner based on evaluative meta-information labelling and indexing.[13–15] The impact of these efforts remains unclear.

More recently, a European project recommended the accreditation of healthcare related software, telemedicine, and internet sites.[16] They suggested a mechanism similar to the marking of electrical goods for software, that national regulatory bodies should be identified for telemedicine, and that a European certification of integrity scheme should be developed for websites. Citing the many impediments to voluntary quality assurance for websites, the authors suggest the development of criteria, modifiable according to the needs of special interest groups, that would be used by accredited agencies to self label conforming websites (not only those offering health information) with a EuroSeal. Monitoring of integrity would be ongoing through cryptographic techniques.

In conclusion, our updated study shows that award systems based on non-validated rating instruments continue to be produced but that most stop functioning soon after their release. Alternative strategies are now flourishing, and whether they are

valid, needed, or sustainable and whether they make a difference is the subject of further research.

What is already known on this topic

The rapid growth of healthcare websites in the 1990s was accompanied by initiatives to rate their quality, including award-like symbols on websites

A systematic review of the reliability and validity of such rating instruments, published in 1998, showed that they were incompletely developed

What this study adds

Few of the rating instruments identified in 1998 remain functional; 51 new instruments were identified

Of the 51 newly identified instruments, 11 were not functional, 35 were available but provided no information, and five provided information but were not validated

Many researchers, organisations, and website developers are exploring alternative ways of helping people to find high quality information on the internet

1 Jadad AR, Gagliardi A. Rating health information on the internet: navigating to knowledge or to Babel? *JAMA* 1998;279:611-4.

2   Moher D, Jadad AR, Nichol G, Penman M, Tugwell P, Walsh S. Assessing the quality of randomized controlled trials: an annotated bibliography. *Control Clin Trials* 1995;16:62-73.
3   Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling and assuring the quality of medical information on the internet. *JAMA* 1997;277:1244-5.
4   Wyatt JC. Measuring quality and impact of the world wide web [commentary]. *BMJ* 1997;314:1879-81.
5   Shon J, Marshall J, Musen MA. The impact of displayed awards on the credibility and retention of web site information. *Proc AMIA Symp* 2000:794-8.
6   Berland GK, Elliott MN, Morales LS, Algazy JI, Kravitz RL, Broder MS, et al. Health information on the internet: accessibility, quality, and readability in English and Spanish. *JAMA* 2001;285:2612-21.
7   Li L, Irvin E, Guzman J, Bombardier C. Surfing for back pain patients: the nature and quality of back pain information on the internet. *Spine* 2001;26:545-7.
8   Suarez-Almazor ME, Kendall CJ, Dorgan M. Surfing the net—information on the world wide web for persons with arthritis: patient empowerment or patient deceit? *J Rheumatol* 2001;28:185-91.
9   Impiccatore P, Pandolfini C, Casella N, Bonati M. Reliability of health information for the public on the world wide web: systemic survey of advice on managing fever in children at home. *BMJ* 1997;314:1875-9.
10 Griffiths KM, Christensen H. Quality of web based information on treatment of depression: cross sectional survey. *BMJ* 2000;321:1511-5.
11 Abbott VP. Web page quality: can we measure it and what do we find? A report of exploratory findings. *J Public Health Med* 2000;22:191-7.
12 Tamm EP, Raval BK, Huynh PT. Evaluation of the quality of self-education mammography material available for patients on the internet. *Acad Radiol* 2000;7:137-41.
13 Eysenbach G, Diepgen TL. Towards quality management of medical information on the internet: evaluation, labelling, and filtering of information. *BMJ* 1998;317:1496-500.
14 Eysenbach G, Diepgen TL. Labeling and filtering of medical information on the internet. *Methods Inf Med* 1999;38:80-8.
15 Price SL, Hersh WR. Filtering web pages for quality indicators: an empirical approach to finding high quality consumer health information on the world wide web. *Proc AMIA Symp* 1999:911-5.
16 Rigby M, Forsstrom J, Roberts R, Wyatt J. Verifying quality and safety in health informatics services. *BMJ* 2001;323:552-6.

How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews

Gunther Eysenbach, Christian Köhler

Abstract

Objectives To describe techniques for retrieval and appraisal used by consumers when they search for health information on the internet.

Design Qualitative study using focus groups, naturalistic observation of consumers searching the world wide web in a usability laboratory, and in-depth interviews.

Participants A total of 21 users of the internet participated in three focus group sessions. 17 participants were given a series of health questions and observed in a usability laboratory setting while retrieving health information from the web; this was followed by in-depth interviews.

Setting Heidelberg, Germany.

Results Although their search technique was often suboptimal, internet users successfully found health information to answer questions in an average of 5 minutes 42 seconds (median 4 minutes 18 seconds) per question. Participants in focus groups said that when assessing the credibility of a website they primarily looked for the source, a professional design, a scientific or official touch, language, and ease of use. However, in the observational study, no participants checked any "about us" sections of websites, disclaimers, or disclosure statements. In the post-search interviews, it emerged that very few participants had noticed and remembered which websites they had retrieved information from.

Conclusions Further observational studies are needed to design and evaluate educational and technological innovations for guiding consumers to high quality health information on the web.

Introduction

Little is known about how consumers retrieve and assess health information on the world wide web. Some surveys have elicited data by using semistructured questionnaires or focus groups,[1–3] but little (if any) unobtrusive observational research has been done to explore how consumers are actually surfing the web. Although several criteria for quality of health websites have been proposed—including disclosure of site owners, authors, and update cycle[4 5]—little or nothing is known about whether and to what degree such markers are recognised or even looked at by consumers or what other credibility markers consumers are looking for. We aimed to obtain qualitative and semiquantitative data to generate some hypotheses on how consumers might search for and appraise health information.

Methods

We used multiple methods of data collection that are commonly used in studies of human-computer interactions,[6] combining focus groups,[7] naturalistic observation of consumers searching the internet, and post-search in-depth interviews. Two researchers independently analysed transcripts by using N5 (NUD*IST 5.0; QSR International, Melbourne) with the grounded theory approach.[8]

Participants in the focus groups and the observational study were mostly healthy volunteers recruited through newspaper advertisements seeking people who had already searched for health information on the web. We selected them on a first come first served basis.

Focus groups

We held three focus group sessions with 6-8 participants each (21 participants in total: five men, 16 women; mean and median age 37, range 19-71 years) in March 2001. Self reported internet experience of the participants ranged from 17 to 84 months (mean 46 months; median 42 months). Each session was facilitated by a skilled moderator and lasted about two

Unit for Cybermedicine and eHealth, Department of Clinical Social Medicine, University of Heidelberg, Bergheimer Str 58, 69115 Heidelberg, Germany
Gunther Eysenbach
senior researcher
Christian Köhler
researcher

Correspondence to:
G Eysenbach
ey@yi.com