

QUALITY CRITERIA MODELS USED TO EVALUATE HEALTH WEBSITES

June Anderson*
Sue McKemmish
Rosetta Manaszewicz

School of Information Management and Systems
Monash University
P O Box 197
Caulfield East Vic 3145
Ph: +61 3 9903 1082, Fax: +61 3 9903 2005
Email: june.anderson@sims.monash.edu.au

Ms June Anderson is a Chief Investigator in the ARC Linkage project – “An Intelligent, User-Sensitive Portal to Breast Cancer Knowledge Online”. In this project she is working on the information identification, assessment and description component as well as the portal design and prototyping component of the project. Information retrieval has been a keen interest of June’s throughout her career as a practitioner for many years in the special library sector, and since 1989 as an academic in the library and information services field at Monash University.

For the last two years June has been the Director of Graduate Studies in the School of Information Management and Systems and has managed a Master’s coursework program with over 450 students. Now she’ll be spending more time on research – an area that she was actively involved in in the early 1990s after being an inaugural VALA Travel Scholar. In her study tour she investigated the Use of Expert Systems Technology in the LIS field in Britain and the USA and reported back on her findings at various conferences.

Ms Roseta Manaszewicz is currently a member of the Breast Cancer Action Group (Vic) Steering Committee, as well as on the Policy Committee of the Breast Cancer Network Australia. As an experience librarian, and consumer representative on several local and national working parties, she brings to the issue of patient information provision, valuable insights into the changing information needs of women and their families across the entire breast care continuum.

Rosetta is working as a Research Fellow at Monash University on the ARC Linkage project – “An Intelligent, User-Sensitive Portal to Breast Cancer Knowledge Online”. In the past Rosetta has also worked on information literacy modules delivered as part of an online education program with Monash. She has also been involved in medical informatics research.

ABSTRACT:

Consumer use and reliance on internet health information is rapidly increasing, along with concerns as to the quality of the information provided in the thousands of health websites currently available. Numerous studies have revealed the inaccuracy of information ranging across many areas, and have highlighted the potential dangers this can present to a lay audience requiring consistent and reliable health information.

This paper provides an overview of the major quality ‘standards’, guidelines and criteria such as DISCERN, Hi-Ethics and the AMA Guidelines that have been developed in recent years to specifically evaluate online health information. It then explores how applicable these schemes are to the information needs of the lay consumer, and how the Monash University led team has developed an innovative metadata schema as part of its Australian Research Council funded project “An intelligent user-sensitive portal to Breast Cancer Knowledge Online” (BCKOnline). Through the use of ‘Quality’ elements and qualifiers in the metadata schema which aim to empower the consumer, the portal prototype will provide users with an assessment of the quality of the resource that can then be applied according to their individual needs and value systems.

Consumer use and reliance on Internet health information has been well documented over the past few years. Estimates as to the actual number of online health seekers varies but, the most often cited statistics reveal that in the United States at least 109 million users have regularly consulted the Internet in order to glean information and/or to manage their health care (Harris Interactive, 2003). European surveys also estimate a large growth in Internet health seekers (Spadaro, 2003). Given the potential for harm which inaccurate and spurious information may hold for the consumer, issues of ‘quality’ information provision and the need for international standards which are capable of alerting the user as to the accuracy, reliability, and authoritativeness of the resource has consequently been the focus of much concern.

Various models, codes of conduct, ‘seals of approval’, guidelines, and accreditation programs have been developed in order to assist the lay consumer to identify online resources which may stake some claim to ‘quality’. The overriding difficulty, however, has been in reaching consensus

as to what constitutes 'quality', what criteria to use to assess it, and how these criteria might be standardized and applied across a multitude of resources relating to a vast range of health conditions. A compounding factor is that information resources are presented in an increasingly diverse array of formats.

The Australian Research Council funded project "An intelligent user-sensitive portal to Breast Cancer Knowledge Online" (BCKOnline) (<http://www.sims.monash.edu.au/research/eirg/bcko.html>) has addressed these issues of quality assurance via an innovative metadata schema, part of which includes a 'quality reporting' element designed to alert users as to the reliability and authority of the resource. Using the BCKOnline portal as a case study, this paper will:

- analyze the content quality criteria used in different models
- analyze user needs in relation to quality information
- provide an overview as to how the BCKOnline research team standardized the findings and translated them into metadata elements and qualifiers that can be used to describe resources and produce reports about their provenance, accuracy, authority and reliability that will enable users to make their own judgements about quality in terms of their individual needs and value systems.

CURRENT MODELS OF QUALITY CRITERIA

Numerous models and 'standards' have been developed which attempt to provide producers of online resources and websites with the standards required to ensure quality and users with the means of evaluating such quality. One study (Kim et. al., 1999) identified over 165 individual criteria covering 12 distinct categories whilst another (Gagliardi and Jadad, 2002) found that in the space of 3 years, 98 complete additional schemes had been devised – although few were still operational. For the purposes of brevity, this paper will concentrate on models which are currently cited in the literature, have some international standing and which were specifically developed, or adapted, for the online environment. The models are:

- Health on the Net – HonCode of Conduct (1997)
- American Medical Association – Guidelines for Medical & Health Information Sites on the Internet (2000)
- Health Internet Ethics – Hi-Ethics (2000)
- European Community – Quality Criteria for health-related websites (2001)
- DISCERN (2000)
- OMNI – Organising medical networked information (1999)
- Internet Health Coalition – Ehealth Code of Ethics (2000)

These models represent the diversity of current thinking on quality assurance. Differences are to be found in their respective purpose and mission, intended audiences, enforcement protocols and emphases on specific evaluation criteria. Generally, the various criteria encompass three major categories – content, privacy/confidentiality issues, and web design or technological

features. This paper will concentrate on the content criteria only, since it is assumed that the search for reliable and authoritative content is generally the catalyst for online health information seeking. Criteria designed to evaluate 'navigability' and other technological features of sites will be addressed in the evaluation phase of the BCKOnline project.

CONTENT QUALITY CRITERIA

An analysis of the above listed models reveals several common threads pertaining to content evaluation, albeit the terminology and emphases differs. Major agreement however is to be found in the 'accountability' criteria which Silberg et. al., (1997) defined as source, attribution, currency, and disclosure. Several of the guidelines also address the issue of 'accuracy' however, the criteria is wide ranging and varies from concordance with published clinical practice guidelines to adherence to 'best practice'.

The following definitions represent a summary of the major qualifiers used in the various schemes and which pertain to content evaluation.

- *Authorship* - this includes both the creator and publisher of the material and in some guidelines the 'authority' of the author. Generally taken to mean credentialing.
- *Accuracy* – what 'evidence' is available to support the claims? Criteria here range from case studies to evidence-based medicine and the obligation to inform users as to which level of evidence has been used.
- *Source* – need for clear citations and/or links to cited evidence.
- *Comprehensiveness* – how extensive is the information given? Does it include areas of disagreement/lack of evidence?
- *Currency* – date of creation, posting, amendment to material(s). This criteria may also apply to the source of information if on another website.
- *Review process* – what is the editorial policy of the site? Has this material undergone peer review, or editorial processing?
- *Purpose* – overall objective of site and material should be evident. Is the aim of the site to sell, promote, or simply 'educational'. The material is then judged according to whether it fulfils its specific purpose.
- *Balance* – this attempts to assess whether material is objective, emotional, biased, or simply promoting one line or product.
- *Language* – whilst most sites argue for 'understandable' information which caters to the specific audience, only Discern specifically endorses the use of the Flesch measure in order to provide content which is directed towards a lay audience.
- *Disclosure* – the documenting of financial, pecuniary interests in the site and its management. This would also include any sponsorship arrangements.

Whilst this 'list' represents the breakdown of major elements considered vital for content quality assessment, few of the listed schemes apply them in a consistent and specified manner. For instance, the source or attribution category may vary from a mandatory published bibliography of sources (HonCode) to the vague injunction of the need for 'reasonable support' (Hi-Ethics). Nor is their intended use uniform. DISCERN is aimed at lay consumers, yet demands a significant

time investment as they work their way through a series of generic questions for each individual resource. Figure 1, on the next page, provides a breakdown of categories and how they apply to individual models.

THE MODELS AND THEIR RELEVANCE TO THE CONSUMER

How applicable are the above schemes and elements to the information needs of the lay consumer? Do they in fact ‘measure what they claim to measure’? (Jadad and Gagliardi, 1998: 614) More importantly, how do such models account for the highly subjective nature of the information seeking process and the impact that motivation, need, and timeliness may have on the user’s perception of quality and how to evaluate it? Does the woman who has just been diagnosed with breast cancer and must decide between a mastectomy or a lumpectomy view the issue of ‘quality’ in the same manner as a healthy 40 year old simply browsing some sites which feature information on mastectomies? The difficulty with many of the current models of quality criteria is that they assume a uniform lay audience, with uniform information needs, and who are generally in need of added protection and/or education. The literature is replete with calls to improve user ‘health literacy’, or the need to ‘educate’ users regarding information ‘accuracy’ and ‘reliability’. As well as assuming a uniform audience, these schemes also assume that users are seeking only the type of knowledge that derives from the biomedical domain and is predominantly text based. There are no specific criteria amongst the previously listed models designed to evaluate the quality of multimedia presentations or bulletin board discussions. The user, in the end, is often confronted with an array of schemes, codes, and guidelines, relevant only to some types of resources, each employing different and confusing terminology, or demanding a significant time investment as with DISCERN. Four models alone have been shown to contain ‘over 100 separate elements that users should read about and understand to be an informed user of a site.’ (Baur and Deering, 2000) To expect any user to deal with such complexity and diversity is both unrealistic and onerous and, possibly what Ferguson calls ‘internet exceptionalism’.

Most (kitemarks or seals of approval) are unstandardised and inaccessible to the public and so are an opaque way of conveying quality to the user....neither system takes into account the range of information available (both on the internet and in other media), the preferences of users or the desirability of going down this route. An alternative is to take a ‘non-exceptionalist’ approach to online health information’.

(Ferguson, 2002:556)

MODEL	AUTHOR	ACCURACY	SOURCE	COMPREHENSIVE	CURRENCY	REVIEW	PURPOSE	DISCLOSURE	BALANCE
Discern	*	*	*		*	*	*		*
AMA Guidelines	*	*	*		*	*		*	
OMNI Gateway	*	*			*		*	*	
eEUROPE Quality Criteria for Health Related Websites	*				*	*		*	
Health on the Net (HonCode)	*	*			*			*	
Hi-Ethics	*	*	*		*	*		*	
Internet health Coalition	*	*		*					*

Figure 1: Models and Breakdown by Category

The BCKOnline prototype embraces this ‘non-exceptionalist’ vision of information provision where ‘quality’ is assessed in the context of the relationship between the service and its usefulness to the intended audience. That is, quality is a component of a relationship between user and resource, rather than an appraisal of the resource alone. Following is a brief outline of this philosophy and the rationale behind the BCKOnline’s approach to determining content quality criteria.

BCKOnline QUALITY CRITERIA: EMPOWERING THE CONSUMER

BCKOnline has as its major objective the development of an ‘intelligent portal’ prototype which will be capable of delivering ‘tailored’ information to the breast cancer community. Through the use of a metadata schema which links user profiles with context-rich resource descriptions, the prototype will endeavour to meet the dynamic and diverse information needs of the user. The user is thus at the centre of the project, rather than the resources themselves. The following assumptions underpin the BCKOnline rationale:

1. The potential audience for the BCKOnline portal is not homogeneous. It represents a diversity of information need, format preference, level of 'readability' requirements, and type of resource desired.
2. 'Knowledge' is contextual – that is, its 'value' to the individual is a dynamic interplay of personal, social, psychological, ethnic, and cultural factors.
3. The perspectives of people with direct and/or personal experience of breast cancer provides a valuable information resource and insight which both complements and enhances the scientific/biomedical view of breast cancer treatment, management and research.
4. Users need to be provided with standardised information about the provenance, authority, accuracy and reliability of the resources they discover through the portal, so they can make their own judgements about their quality with reference to their needs, purposes and value systems.

Currently the Australian Government Locator Service (AGLS) does not provide a 'quality' metadata element. One 'value-added' feature of the prototype will be the provision of information relating to the quality of a resource through the use of metadata elements and qualifiers about its provenance, authority, accuracy and reliability. The elements, qualifiers and the encoding schemes which control their data values have been derived from the conceptual framework developed to analyse existing quality codes, presented earlier in the article, and an analysis of user needs relating to quality. Information relating to the quality of the resources will focus on their content and provenance, not on the aesthetics of web design, technological features or confidentiality and privacy issues. Users of the portal will be able to view a report about the quality of the resources that match their search. In relation to some "quality attributes", the listing of resources resulting from an initial search will be "signposted", e.g. whether or not they contain references and citations of sources. The quality report will be derived from metadata elements and qualifiers that describe the credentials of those involved in the production and delivery of the resource - the creators, contributors and publishers, and publication review or refereeing processes that apply to the resource, as well as attributes of the resource itself, e.g. its evidence base, purpose, balance and currency. It will provide a basis on which individual judgements about quality can be made. The user will be free to view this data, or simply to ignore it. The quality concepts that will be encoded via the metadata schema, their rationale, potential usefulness and relationship to existing AGLS elements are discussed below (Figure 2).

Credentials Of Resource Producers

Individual names of authors, publishers and other contributors to the intellectual content of a resource, whether they be clinicians or organizations, are unlikely to be useful yardsticks for the majority of users. How many consumers would be aware of the standing of the multitude of organizations scattered throughout the world? Hence defining authorship or publisher through such categories as 'lay author', 'clinician', 'researcher', 'consumer group', 'commercial organization' or 'educational institution' was felt to be far more helpful. This also facilitates the exercise of personal choice and prioritizing. Some users may place greater credence in material which originates from consumer groups for example, than material provided by governments.

'QUALITY' ELEMENTS/REFINEMENTS	USED TO PROVIDE INFORMATION ON	ENCODING SCHEME/QUALIFIERS
AGLS ELEMENTS: CREATOR/PUBLISHER/ CONTRIBUTOR BCKO REFINEMENT : SECTOR	CREDENTIALS	Lay author(s) Clinician(s) Researcher(s) Consumer Group Commercial body/group Educational institution Government body Medical organization Cancer organization
AGLS ELEMENT: DATE REFINEMENTS	CURRENCY	Date created Date modified Date posted
BCKO 'QUALITY' ELEMENT REVIEW PROCESS	AUTHORITY	Editorial board Peer review process No review/editorial process
BCKO 'QUALITY' ELEMENT REFERENCES	ATTRIBUTION	Contains references Does not contain references
BCKO 'QUALITY' ELEMENT EVIDENCE-BASED	'ACCURACY'	Meta-analysis Randomized clinical trial Case/cohort study Review Consensus opinion Personal opinion
BCKO 'QUALITY' ELEMENT PURPOSE	PURPOSE	Educational/informative Commercial Reportage of results Discussion forum
BCKO 'QUALITY' ELEMENT BALANCE	BALANCE	Controversial issue – noted Controversial issue – not noted Non-controversial issue

Figure 2: Quality concepts encoded via the Metadata schema

Review Process

Given the profusion of available resources, the consumer and even the medical professional has no real means of assessing relative quality. How does one judge the respective merit of the Australian Medical Journal, as opposed to the Journal of the American Medical Association? In this context the presence of a reviewing or refereeing process, or editorial board, might be seen

as a useful criterion to inform choice. It is again the user who will assign 'value' according to her individual predilections. For instance, a user may give greater credence to 'peer review' than a 'review board'. The quality report will provide information about this.

Citation of Sources

Some users may want to be reassured of the authoritativeness of a resource through the evidence provided in endnotes or citation of sources in references. Since references are not consistently provided in all health materials, users will be advised whether or not the resource includes references.

Evidence-based

Few users would be aware of the NHMRC breakdown of 'levels of evidence'. Nor would the categorization of information according to guidelines be consistent given the global diversity of standards and/or recommendations. The issue of mammographic screening for women under fifty for example illustrates the range of possible recommendations across countries. Furthermore, many resources of relevance to users, such as experiential knowledge and or complementary medicine, do not fall into the category of scientific or medical material of the kind that can be quality rated according to evidence-based criteria. To overcome such inconsistency, encompass different kinds of resources, and to ensure greater clarity for the user it was determined that resources would be assessed according to the nature or type of content under consideration. It was decided to categorize individual resources according to the following scheme: 'personal opinion', 'review material', 'case study', 'randomized trial', and so on, and then to apply appropriate criteria/standards to each category. Such a classification scheme also enables the categorization of experiential and anecdotal documents without necessarily privileging one form of knowledge over the other.

Purpose

BCKOnline recognizes the possibility that information emanating from a commercial website may be educational and/or informative, whilst conversely, resources from a consumer group could quite feasibly be endorsing a commercial product. Hence, the aim is to inform users whether individual resources themselves are 'educational' or 'commercial'. In this manner the 'purpose' qualifier would be at the individual resource level, rather than the domain level, hence avoiding some of the confusion which listings such as 'org' and 'com' may have in terms of their affiliations and subsequent purpose.

Balance

Information will also be provided to the user on whether the subject matter of the resource is controversial, e.g. concerned with an issue where opinions vary as to the efficacy of treatment, or the nature of cause. In breast cancer, as with most health conditions, certain areas have not attained medical consensus. The quality report will indicate to the user if this is the case in relation to the resource they are viewing.

Currency

Through the qualifiers of date created, modified, and posted, as well as the 'update' qualifier, it is intended to provide the user with information relating to the currency of the resource.

Conclusion

Evaluating the accuracy, provenance, authority and reliability of web-based health information is increasingly seen as a major priority given the volume of information available and its sometimes spurious or fraudulent nature. The various quality models described in this paper all attempt to ensure the validity and reliability of online data through a variety of protocols, and standards. The difficulty, however, with the vast majority of these models is the lack of consistency and the often complex and onerous burden they may place on the consumer attempting to follow and understand their criteria. Nor do these models address the contextual and subjective needs and values of the lay consumer, nor the range of resource types which she may deem relevant and important. The BCKO research project, through its innovative development of metadata elements and qualifiers that will enable the portal to provide users with a report on factors relevant to an assessment of the quality of the resource in terms of user needs and value systems, attempts to address such issues. It does so through a quality reporting and 'signposting' system applicable to a wide variety of information resource types and target audiences. In this manner the issues of personal autonomy and decision-making capacity are respected and enhanced.

REFERENCES:

Baur, C., and Deering, MJ. (2000.) "Proposed frameworks to improve the quality of health web sites: Review". *Medscape General Medicine*, 2(3). <http://www.medscape.com>

Ferguson, T. (2002) "Against internet exceptionalism: there's nothing radically different about information on the web". *British Medical Journal*, 324: 556-557.

Gagliardi, A., and Jadad, A. (2002) "Examination of instruments used to rate quality health information on the internet: Chronicle of a voyage with an unclear destination". *British Medical Journal*, 324: 569-573.

Harris Interactive. (2003) "No significant change in the numbers of 'cyberchondriacs' – Those who go online for health care information". *Health Care News*, 3 (4) March 28. http://www.harrisinteractive.com/news/newsletters/healthnews/HI_HealthCareNews2003Vol3_Is04.pdf

Health on the Net Foundation. (1997) *Operational definition of the HONcode principles*. <http://www.hon.ch/HONcode/Guidelines/guidelines.html>

Hi-Ethics. (2000) *Health internet ethics: Ethical principles for offering internet health services to consumers*. <http://www.hiethics.org/Principles/Index.asp>

Jadad, A., and Gagliardi, A. (1998) "Rating health information on the internet: Navigating to knowledge or to babel?" *JAMA: Journal of the American Medical Association*, 279(8): 611-614.

Kim, P., Eng, T., Deering, MJ., and Maxfield, A.(1999) "Published criteria for evaluating health related web sites: Review". *British Medical Journal*, 318: 647-649.

Silberg, W., Lundberg, G., and Musacchio, R. (1997) "Assessing, controlling, and assuring the quality of medical information on the internet: Caveant lector et viewer – Let the reader and viewer beware". *JAMA: Journal of the American Medical Association*, 277(15); 1244-1245.

Spardaro, R. (2003) *European Union citizens and sources of information about health*. Brussels, EORG: The European Opinion Research Group. http://europa.eu.int/comm/health/ph_information/documents/eb_58_en.pdf