GPHE 133000—13/9/2005——166233

Pharmacy Education, 2005; 5(3): 1–6

One step forward, two steps back: The World Wide Web as an information resource

MICHAEL A. VERONINA

Department of Pharmacy Practice, Texas Tech University Health Sciences Center, School of Pharmacy, 1300 S. Coulter Road, Room 404, Amarillo, TX 79106, USA

Abstract

The World Wide Web offers the potential for comprehensive shared access of reliable and valid health information and educational resources; hence a major step towards enriching pharmacy education. However, educators and students must now consider web site information that may change, move, expire and disappear with no record. Also, many web sites lack key indicators of quality such as the source of the material presented, authorship, or how current the information is. In a sense, these issues may be viewed as major "steps backward" toward using web-based information to enrich education. Although no comprehensive, organized strategy exists to ensure high quality and availability of web sites at all times, several initiatives exist to assist viewers in dealing with web site disappearance and evaluating the quality of information presented. Awareness of these issues is necessary for the World Wide Web to be useful as a valid information and educational resource.

Keywords: *World Wide Web, internet reference, quality initiatives, web site quality, web site access, URL disappearance*

Introduction

For students and scholars, the World Wide Web is proving to be a valuable medium to obtain information on a variety of topics, and can enrich both traditional and distance courses.

It is evident, though, that this technological advancement is not without shortcomings. Compared to traditional print materials, web-based information is proving to be unreliable in terms of accessibility and quality. In some respects this can be viewed as "a step backward" in terms of educational advancement. The purpose of this paper is to provide a brief overview of the quality and accessibility issues that students and educators may face with increased use of the World Wide Web as an information resource.

One step back: Unreliable access

In the medical, scientific and educational literature, an alternative to the printed page is an image or "digital object" known as a web page (Lyman, 2005). Web-based information is often referenced and cited in the form of a URL address. A check of the major journals such as *Science* or the *New England Journal of Medicine* often shows at least one URL web site address as a reference citation listed at the end of a report. URL referencing may be even more pervasive depending on the journal and the topic of the report. For instance, a research report on internet pharmacies that appeared in the *European Journal of Clinical Pharmacology* had 32 total reference citations, of which 23 (67%) of these were URL Web site addresses (Bessell, Silagy, Anderson, Hiller & Sansom, 2002).

Known as "dead link", this term refers to a web reference such as a URL address or link that does not accurately point to an existing web page (Spinellis, 2003). A variety of reasons exist for these nonexistent links (Veronin, 2003), but in any event, this leaves the reader with no concrete evidence for a reference or citation. Educators and students must now consider web site information that may be created, changed, moved, expired and disappear, with no

Correspondence: Michael A. Veronina, Department of Pharmacy Practice, Texas Tech University Health Sciences Center, School of Pharmacy, 1300 S. Coulter Road, Room 404, Amarillo, TX 79106, USA. Tel: 806 356 4015x319. Fax: 806 356 4034. E-mail: michael.veronin@ttuhsc.edu

ISSN 1560-2214 print/ISSN 1477-2701 online © 2005 Taylor & Francis
DOI: 10.1080/15602210500330175

record of the information being preserved (Veronin, 2002).

Research progresses by knowing what has been done previously, and when new scholarly work is produced, it is imperative that information sources are cited accurately and in detail.

How prevalent is the problem? Quite prevalent, as several studies have drawn attention to the problem of URL disappearance and loss of access to web citations, some of which are discussed below.

In MEDLINE, the US National Library of Medicine's bibliographic database, one major study reported that from 1997 to 2002 the average annual increase in the number of abstracts containing a URL was 47%, compared with a consistent average annual increase of 5% in the total number of MEDLINE records over the same period (Wren, 2004). The same study revealed that nearly 19% of the Web sites in abstracts on MEDLINE, were consistently unavailable.

Recently, a report from the University of Colorado in Denver found that about 12% of the internet addresses cited in three high impact periodicals, the *New England Journal of Medicine*, *Journal of the American Medical Association*, and *Science* were unavailable two years after publication (Dellavalle et al., 2003).

In an earlier report, the author examined 184 Web sites available in 1999 with information related to the herbal remedy Opuntia. Approximately three years after initial posting, over two-thirds of the health-related web sites reviewed could not be found or had moved with no forwarding URL address (Veronin, 2002).

Some promising initiatives are being developed to assist Web users in dealing with Web site loss, and are summarized in Table I. The approaches differ, but the goal is essentially the same: to provide reliable access to web-based references.

Another step back: Unreliable Web site quality

Regarding scientific research, Moher et al. (1995) summed up the importance of quality this way: "Quality gives us an estimate of the likelihood that the

Table I. Some initiatives that address Internet reference disappearance.

Initiative	Sponsor	Description
Internet archive	San Francisco-based public nonprofit organization	A digital library of web pages created with the goal of cataloging all of the past and present publicly available material on the World Wide Web. The Wayback Machine feature allows users to view previous versions of web sites or those that no longer exist.
National digital information infrastructure and preservation program (NDIIPP)	U.S. Library of Congress	A collaborative public service project designed to preserve digital information. Areas addressed include web sites, electronic journals, electronic books, digitally recorded sound, digital moving images and digital television.
Google virtual library	Google, Inc. in collaboration with Oxford University and research libraries	Library holdings will be converted into digital files freely searchable over the web. Digital card catalog and searchable library will be created for books, scholarly papers and special collections.
Open archival information system (OAIS)	RLG nonprofit library consortium; online computer library center (OCLC), and others	Conceptual framework for preservation of digital information. A reference model used by several organizations for planning and designing digital repositories.
Digital object identifier (DOI) system	International DOI Foundation	A unique alphanumeric string assigned to a digital object such as an electronic journal article. Provides a means to consistently identify the object on a network, such as the www.
CrossRef	Nonprofit network of publishers	Infrastructure for linking citations across publishers for online publications. Based on the DOI system, facilitates reliable reference linking throughout online scholarly literature.
LOCKSS	Stanford university uibraries	Stands for "Lots of Copies Keep Stuff Safe." Open source software that creates inexpensive accessible copies of e-journal content. Allows librarians to preserve authorized content they purchase.
Uniform resource name (URN)	Internet engineering task force (IETF)	Similar in format and function to URL, except resource is not subject to disappearance when digital content is moved.
Persistent uniform resource locator (PURL)	Online computer library center (OCLC) research	A special type of URL designed to enhance reliability of Web site access. A PURL makes use of an intermediate resolution service to return the information resource (URL) to the client.

results are a valid estimate of the truth". A similar level of importance should exist for information on the World Wide Web.

Quality of web documents has been an issue since the early days of the web. According to Lindberg & Humphreys (1998), many web pages "resemble books published in the 15th century and audiovisual materials produced a few decades ago".

Today, the web is a strange mix of information content from meaningless pages to esoteric online journals. Even with recent advances in web technologies, many health-related sites still lack key information such as the source of the material presented, who authored it, or how current the information is. Because anyone with basic computer skills can publish on the web, the veracity and validity of information that the viewer is provided must be considered. Many health-related web sites may be "professional" in appearance, yet may offer little benefit, and may even be harmful (Walji et al., 2004). The potential for misleading and inaccurate health information on the web has been pointed out repeatedly by professional groups and individual authors and organizations (Delamothe, 2000).

In the academic world, evidence-based clinical and scientific reports that have gone through the peer review process provide confidence to the reader for the information's validity. What assurances do educators have regarding information presented on the World Wide Web? Ideally, sites would indicate whether the information provided has been subjected to outside authoritative review, but in reality, information on quality assurance is rarely disclosed (Bernstam, Shelton, Walji & Meric-Bernstam, 2005).

No absolute consensus exists as to what attributes a health-related Web site should possess to be deemed "high quality", and hence containing valid and accurate information. Distinguishing characteristics necessary to produce high quality web-based health information have been described by several authors (Risk and Dzenowagis, 2001; Mitretek Systems Health Information Technology Institute, 2000). Ideally, if a web site possesses attributes delineated by these standards of quality, a web site viewer can

Table II. Elements of a Web Site that affect quality.

Attribute	Definition/Comments
Source	Refers to the individual or organization responsible for information presented on the Web site, and should be apparent to the viewer. "Consider the source" is the mantra for evaluating credibility of information in all forms of media, including the www.
Qualifications and credentials	Web sites often promote the qualifications and/or credentials of their organization and authors. If these elements are missing or not readily apparent, the credibility of the source of the information should be regarded as suspect.
Conflict of interest	Web site viewers should be aware of potential sources of bias with the information provided. Unlike the peer-review system of medical journals, Web site authors may not disclose potential conflicts of interest, such as sources of funding.
Currency	Currency means that information is up-to-date with regard to the present state of medical, scientific and clinical knowledge. The date that the original document was created subsequent content was posted should be displayed on the site's main page.
Purpose	Commonly implied or stated within the site's homepage, the reason for creating the site or intent of the developer should be apparent. Often in the form of a "mission statement," the site's content should be appropriate to this purpose or mission.
Relevance	A relevant article provides information that is topical and useful. Relevant documents "must either give the user new knowledge, correct old knowledge, or update existing knowledge." The same distinctions can be applied to pages retrieved from the Web.
Accuracy	Accurate Web content is verifiable and evidence-based. Ideally, the site should identify the data from which conclusions are based. Scientific and clinical evidence cited to support a viewpoint should be clearly presented.
Context	Relates to a site's purpose, the setting for information affects the interpretation and meaning. For example, identical facts about vaccines would be considered differently in the context of disease prevention versus bioterrorism.
Scope/Totality of the topic	Web site viewers should decide if information is reasonably thorough on a given topic, to avoid coming to erroneous conclusions. That is, web page authors may omit critical information, or put forth slanted views, possibly driven by a hidden agenda.
Disclaimer	The general definition of disclaimer is "a denial or disavowal of legal claim." Ideally, a web site disclaimer would call attention to limitations of the information, such as emphasizing that the content is general health information, not medical advice.
Design	To be effective, information should be displayed in a format that is accessible, logical, and meaningful to the user. Design relates to an understanding of the tasks the viewer seeks to accomplish with the desired information.
Links	Linking creates relationships among seemingly disparate information. There should be a perceived order by which the viewer navigates within or out of the site. Effective design of links may determine the success or failure of the delivery of information.

have confidence that the information contained within the site is credible, and in essence, trustworthy. Major attributes contributing to web site quality, though not exhaustive, are listed in Table II.

Assessment of Web site quality

Different approaches exist for rating the quality of web sites and several organizations have developed instruments for rating web site quality (Risk & Dzenowagis, 2001). Often there is no clear delineation of the mechanisms by which quality is determined nor is it known how widely they are being implemented (Wilson, 2002; Bernstam et al., 2005). Many are designed to assist the viewer in recognizing web site quality (and presumably quality information), and some assist web site developers in creating quality sites. Major initiatives for rating health information quality are described in Table III.

Discussion

In a 1996 speech, former President Bill Clinton remarked, "When I took office, only high energy physicists had ever heard of what is called the World Wide Web … now even my cat has its own page" (White House Press Release, 1996). This statement underscores the fact that web sites and internet references have become routine, and even may be expected to increase as 7 million pages of new information, appears daily on the World Wide Web (Dellavalle et al., 2003).

Disappearance of web references and unreliable information quality impose barriers to an educator's goal of comprehensive shared access to reliable and valid health information. Currently, although several suggestions are offered, no comprehensive, organized strategy exists to ensure high quality and availability of web sites at all times (Whitfield, 2004).

Web-based materials can enrich education by making information rapidly accessible and easy to update. Versatility is enhanced by allowing access to content not suitable for print media, such as high-resolution images, motion video, animations, simulations, and program source code (Dellavalle et al., 2003).

At this time, since no federal or state laws regulate health-related web sites (Charatan, 2002). With regard to judging quality, responsibility still lies largely with web viewers.

Some of the criteria for assessing quality are rather basic, which the viewer can apply with minimal effort. For example, when viewing a web site, attention should be given to the site's owners or sponsors, and the motive behind creation of the site to expose potential conflicts of interest. Author names and credentials, and references to information, and dates of posted information should be apparent.

With regard to web site disappearance, it is clear that educators should be aware of some caveats in selecting links or addresses to their web-based materials. Web sites are not maintained or are abandoned, computer

Table III. Approaches to assist viewers assess Web site quality.

Evaluation approach	Example(s)	Comments
Codes of conduct	HON Code. Health on the net foundation code of conduct; American Medical Association (AMA)	Provide recommendations for the content and development of web sites, based primarily on principles of ethical behavior. Quality criteria (Table I) may be contained in codes of conduct.
Symbol/Logo	Health internet ethics (Hi-Ethics); HON Code logo	Much like the Good Housekeeping "seal of approval," web sites may display a quality label (also known as a "quality symbol" or "quality logo") that signifies adherence to a code of conduct by a web site developer.
Third-party accreditation	MedCERTAIN; Organized medical networked information (OMNI)	A variation on the quality symbol system; an independent third party awards "site accreditation." The organization or individual who develops the web site must submit documentation of compliance with quality standards delineated by the impartial third party.
Viewer questionnaires	DISCERN; Health Summit Working Group (HSWG) IQ Tool	Also known as "user guidance systems," this manner of assessing web site quality relies on a set of questions to be answered by viewers to validate information. Viewer may access the series of questions from a link or symbol displayed on the site.
Web browser filters	SmartFilter™; Websense™	Web-filtering software can prevent a computer's browser from displaying specific information on web sites. Sites with undesired information may be accepted or rejected based on predefined criteria.
Gateway catalog	Catalog and index of french-speaking medical sites (CISMeF) project	Viewers access specific web sites where information has been cataloged by an independent third party. Medical experts critique medical information resources based on quality and relevance to a particular audience. Electronic descriptive "tags" are place on the information, then classified and the descriptions are stored in a database.

servers close down, homepages disappear when researchers move, and files are capriciously redirected or renamed (Lawrence et al., 2001). The web is not an organized library; it has taken many years for libraries to evolve to provide reliable archival services for books, peer-reviewed periodicals, and other printed publications. Until reliable methods are developed, access and assessment of quality health-related information on the web remains a challenging undertaking for educators.

Acknowledgements

The author received no special funding in the preparation of this manuscript.

References

Bernstam, E. V., Shelton, D. M., Walji, M., & Meric-Bernstam, F. (2005). Instruments to assess the quality of health information on the World Wide Web: What can our patients actually use? *International Journal of Medical Informatics*, *74*, 13–19.

Bessell, T. L., Silagy, C. A., Anderson, J. N., Hiller, J. E., & Sansom, L. N. (2002). Quality of global e-pharmacies: Can we safeguard consumers? *European Journal of Clinical Pharmacology*, *58*, 567–572.

Charatan, F. (2002). "Buyer beware" remains US policy towards information on the net. *BMJ*, *324*, 566.

Delamothe, T. (2000). Quality of websites: Kitemarking the west wind. Rating the quality of medical websites may be impossible. *BMJ*, *321*, 843–844.

Dellavalle, R. P., Hester, E. J., Heilig, L. F., Drake, A. L., Kuntzman, J. W., Graber, M., & Schilling, L. M. (2003). Information science. Going, going, gone: Lost Internet references. *Science*, *302*, 787–788.

Lawrence, S., Coetzee, F., Glover, E., Pennock, D., Flake, G., Nielsen, F., Krovetz, R., Kruger, A., & Giles, C. L. (2001). Persistence of web references in scientific research. *IEEE Computation*, *34*, 26–31.

Lindberg, D. A., & Humphreys, B. L. (1998). Medicine and health on the internet. The good, the bad, and the ugly. *The Journal of the American Medical Association*, *280*, 1303–1304.

Lyman, P. Archiving the World Wide Web. Problem statement: Why archive the Web?, Library of Congress. Digital Preservation, National Digital Information Infrastructure and Preservation Program. From the Internet: http://www.digitalpreservation.gov. Accessed 05/2005.

Mitretek Systems Health Information Technology Institute. Criteria for assessing the quality of health information on the internet-policy paper. From the Internet: http://www.hitiweb.mitretek.org/docs/policy.html. Accessed 05/2005.

Moher, D., Jadad, A., Nichol, G., Penman, M., Tugwell, P., & Walsh, S. (1995). Assessing the quality of randomized controlled trials: An annotated bibliography of scales and checklists. *Controlled Clinical Trials*, *16*, 62–73.

Risk, A., & Dzenowagis, J. (2001). Review of internet health information quality initiatives. *Journal of Medical Internet Research*, *3*, e28.

Spinellis, D. (2003). The decay and failures of web references. *Commun ACM*, *46*, 71–77.

The Quotations Page, From the Internet: http://www.quotationspage.com/quote/24883.html. Accessed 05/2005.

The White House. Office of the Press Secretary (Knoxville, Tennessee.) October 10, 1996, Excerpts from Transcribed Remarks by the President and the Vice-President to the people of Knoxville on the Internet for Schools. From the Internet: http://www.govinfo.library.unt.edu/npr/library/speeches/101096.html. Accessed 04/2005.

Veronin, M. A. (2002). Where are they now? A case study of health-related web site attrition. *Journal of Medical Internet Research*, *4*, e10.

Veronin, M. A. (2003). 'URL 404 file not found': Dealing with the transient nature of the web. *The Journal of Audiovisual Media in Medicine*, *26*, 153–155.

Walji, M., Sagaram, S., Sagaram, D., Meric-Bernstam, F., Johnson, C., Mirza, N. Q., & Bernstam, E. V. (2004). Efficacy of quality criteria to identify potentially harmful information: A cross-sectional survey of complementary and alternative medicine web sites. *Journal of Medical Internet Research*, *6*, e21.

Whitfield, J. (2004). News. Web links leave abstracts going nowhere. *Nature*, *428*, 592.

Wilson, P. (2002). How to find the good and avoid the bad or ugly: A short guide to tools for rating quality of health information on the internet. *BMJ*, *234*, 598–602.

Wren, J. (2004). 404 not found: the stability and persistence of URLS published in MEDLINE. *Bioinformatics*, *20*, 668–672.

References: Table I

Crossref Web site. From the Internet: http://www.crossref.org. Accessed 05/2005.

The digital object identifier system. From the Internet: http://www.doi.org. Accessed 05/2005.

Editors, the Internet archive. Universal access to human knowledge. Frequently asked questions. From the Internet: http://www.archive.org/about/faqs.php. Accessed 05/2005.

Library of Congress National Digital Information Infrastructure and Preservation Program. Digital preservation. From the Internet: http://www.digitalpreservation.gov. Accessed 05/2005.

Markoff, J. & Wyatt, E., Technology; Google is adding major libraries to its database. New York Times. Business/Financial Desk. 2004 December 14; Sect A, Page 1, Column 6. From the Internet: http://www.nytimes.com/2004/12/14/technology/14google.html. Accessed 05/2005.

The PURL home page. From the Internet: http://www.purl.oclc.org. Accessed 05/2005.

Reich, V. & Rosenthal, D., (2004). Preserving today's scientific record for tomorrow. *BMJ*, *328*, 61–62.

RLG Web site. OAIS activities. From the Internet: http://www.rlg.org/longterm/oais.html. Accessed 05/2005.

Spinellis, D. (2003). The decay and failures of Web references. *Commun ACM*, *46*, 71–77.

References: Table II

Eysenbach, G., Powell, J., Kuss, O., & Eun-Ryoung, S. (2002). Empirical studies assessing the quality of health information for consumers on the World Wide Web. A systematic review. *JAMA*, *287*, 2691–2700.

Hersh, W. (1994). Relevance and retrieval evaluation: perspectives from medicine. *J Am Soc Inf Sci*, *45*, 201–206.

Mitretek Systems Health Information Technology Institute (2000). Criteria for assessing the quality of health information on the Internet-policy paper. From the Internet: http://www.hitiweb.mitretek.org/docs/policy.html. Accessed 05/2005.

Risk, A. (2002). Commentary: on the way to quality. *BMJ*, *324*, 601.

Webster's New Collegiate Dictionary (1976). Springfield, MA: G & C Merriam Company, p 325.

Web Style Guide. 2nd ed. From the Internet: http://www.webstyleguide.com. Accessed 05/2005.

References: Table III

Health on the Net Foundation. Principles. From the Internet: http://www.hon.ch/HONcode/Conduct.html. Accessed 05/2005.

Richardson, C. R., Resnick, P. J., Hansen, D. L, Derry, H. A. & Ridehout, V. J. (2002). Does pornography-blocking software block access to health information on the Internet? *JAMA*, *288*, 2887–2894.

Risk, A., Dzenowagis, J. (2001). Review of Internet health information quality initiatives. *J Med Internet Res*, *3*, e28.

Robison, L. (2001). In with the good, out with the bad. Use filtering software to control spam and Web content. How the Internet Works. Smart Computing Reference Series. Part 1, *5*, 181–184.

Soualmia, L. F., Darmoni, S. J. (2005). Combining different standards and different approaches for health information retrieval in a quality-controlled gateway. *Int J Med Inf*, *74*, 141–150.

Wilson, P. (2002). How to find the good and avoid the bad or ugly: a short guide to tools for rating quality of health information on the Internet. *BMJ*, *234*, 598–602.

Winker, M. A., Flanagin, A., Chi-Lum, B., White, J., Andrews, K., Kennett, R. L., DeAngelis, C. D., & Musacchio, R. A. (2000). Guidelines for medical and health information sites on the Internet. Principles governing AMA Web sites. *JAMA*, *283*, 1600–1606.

Author Queries

JOB NUMBER: MS 133000

JOURNAL: GPHE

Q1 The Quotations Page is given in the reference list but not cited in the text. Please check.