The New Navigators: from Professionals to Patients
R. Baud et al. (Eds.)
IOS Press, 2003

Diabetes Websites Accredited by the Health On the Net Foundation Code of Conduct: Readable or Not?

Sanja Kusec[a], Ognjen Brborovic[a], Dean Schillinger[b]

[a]*Andrija Stampar School of Public Health, Medical School, University of Zagreb, Croatia*
[b]*University of California San Francisco, Primary Care Research Center, Department of Medicine, San Francisco General Hospital, San Francisco, CA, USA*

Abstract

Health information posted on the Internet has become a popular mode of communication with the population at large because millions of people now use the Internet to gather health information. Many studies on readability have shown that patient education information is frequently written at the reading level too high for the average population to understand, and the same holds true for health information on the Internet. The aim of this study was to determine the readability levels of health information found on diabetes-related websites displaying HONcode logo, which indicates to high quality of the information provided. The 99 websites tested for readability using the Flesch Reading Ease formula and Flesch-Kincaid level showed FRE score 2.1 to 79.6, with the mean 41.7 (10^{th} grade, 8^{th} month Flesch-Kincaid level), which indicates that 86.9% of these materials would be too difficult to read for the average adult population. It is suggested that the readability level, and the name of the formula used, be used on the websites themselves to help the Internet users decide which sites could be of the greatest benefit to them.

Keywords

Readability; Health On the Net Foundation; Diabetes

1. Introduction

Doctor-patient interaction has been changing from a paternalistic to a more cooperative approach. In the context of chronic disease care, a critical goal of the clinical intervention is to motivate the patient to become a partner in their care and to convey the self-management skills that make them more effective partners. A significant portion of office-based physician-based communication is focused on assessing and encouraging self-care. The Internet has the potential to extend this dialogue beyond the office walls by both better preparing patients for their limited time with their physician and by providing reinforcement of key concepts and support for appropriate self-care at home.

Health information designed for patients and posted on the Internet has become a popular mode of communication with the population at large. Nearly half of US residents and one-third of Europeans use the Internet for health purposes [1]. The Internet thus has great potential as a resource for consumers - in theory, providing an accessible forum to share, disseminate, and use health information [2]. With so many people searching the Internet for health and medical information, the quality of that information becomes critical. Recently, the reliability and accuracy of health-related websites has been called into question [2].

There has been a need to both evaluate health information on the Internet and empower consumers as to how to assess quality of the health information. A number of authors and organizations have developed a set of criteria or guidelines that include indicators intended to help Internet users determine the reliability of health information on the Internet [2].

The Health On the Net Foundation Code of Conduct (HONcode) for medical and health websites addresses one of the Internet's main healthcare issues: the reliability and credibility of information [3]. Health On the Net Code of Conduct was launched in early 1996 with the aim of raising the quality of healthcare information available 'on the Net'. Being a self-regulatory, voluntary certification system based on an 'active seal' concept, it is primarily intended for healthcare site developers and publishers but also helps users identify sources of reliable information. It tends to standardize the reliability of information by defining a set of rules based on basic ethical standards in the presentation of information [3]. Today, the HONcode seal is seen on thousands of websites that abide by eight principles: authority, complementarity, confidentiality, attribution, justifiability, transparency of authorship, transparency of sponsorship, and honesty in advertising and editorial policy [3].

The HONcode approval has been shown to be one indicator of website quality that is recognizable by lay people. Websites with a HONcode logo are almost four times more likely to be displayed on a more accurate site than on a less accurate site [2]. Since the HONcode logo appears to be an indicator of accuracy and reliability of health websites, and since it is the most well-established means of demonstrating accreditation, we examined whether HONcode approval was associated with quality across an additional criterion not included in the current eight HONcode principles: readability.

Readability, one measure of the language comprehensibility of text, is often used to assist writers, editors, teachers, and librarians in matching the difficulty of written material with the reading ability of the intended readers: a good match improves communication and learning [4]. However, many studies have shown that most patient literature written in English fails to conform with the current standards of readability [5], and the same holds true for information found on the Internet [6-10].

Many recent studies on readability have examined chronic disease educational materials on such conditions as diabetes mellitus, asthma and cancer [11]. According to the HON statistics, the most frequently searched terms by MedHunt in 2001 and 2002 were: 'diabetes', 'asthma', and 'cancer' [3]. The aim of this study was to determine the reading levels of health information found on diabetes-related HONcode accredited websites, and to propose additional steps for raising the quality and accessibility of health information on the Internet.

2. Methods and Material

We identified HONcode accredited diabetes-related websites entering the word *'diabetes'*, and using the HONcodeHunt search engine (on October 30, 2002). We then limited the search to the English language. A physician experienced in diabetes care and web-based research then reviewed sites for their direct relevance to diabetes and excluded those that were unretrievable, required registration, contained only links to other sites, search engines or redirection to other sites, or sites containing information of less than 100 words. We then sampled totals of 100-200 word text (continuous text

without graphics) from each website by identifying text from the beginning of the site and text a similar length from the end of the website's home page or direct page on diabetes. The second part of the sample text was used to avoid the nontypical writing style usually observed in introductory paragraphs [12].

We then applied the 'copy/paste' procedure to copy sample texts into a Microsoft Word document, and used 'Tools' menu, under 'Spelling and Grammar', to obtain readability statistics (FRE and Flesch-Kincaid scores) for each website. The Flesch Reading Ease (FRE) score and Flesch-Kincaid reading level are two of the most widely used systems for scoring readability [10]. The FRE score ranges from 0 (most difficult to read) to 100 (most easy to read), with ideal score of 60-70 (called 'standard') [12], whereas Flesch-Kincaid score converts the Flesch scale into a grade level estimate. Both scores have been in use for more than 40 years and their validity has been well demonstrated [10]. Briefly, these systems are based on average sentence length in words and number of syllables per 100 words to estimate text difficulty [12].

We then categorized websites according to the location and type of website developer as determined in HON short reports, available upon HONcodeHunt search, for each HONcode accredited website.

3. Results

We identified 293 diabetes-related HONcode accredited websites using HONcodeHunt search. Limitation to English language and exclusion criteria excluded 194 websites. The remaining 99 websites were tested for readability using Flesch Reading Ease and Flesch-Kincaid scores. Expressed by FRE scores, the obtained readability levels ranged from 2.1 to 79.6, with the mean 41.7, and median 40.8; expressed by Flesch-Kincaid scores, the readability levels ranged from 4.7 (corresponding to 4th grade, 7th month reading level) to 12.0 (12th grade), with the mean 10.8 (10th grade, 8th month), and median 12.0 (12th grade). The number of studied websites by FRE scores is shown in Table 1.

Table 1 Readability levels of 99 diabetes-related HONcode accredited websites in English language as expressed by Flesch Reading Ease scores

FRE scores*	Difficulty	Example texts	No of Web sites (%) (n=99)
0-30	Very difficult	Scientific journals	21 (21.2%)
31-50	Difficult	Academic journals	51 (51.5%)
51-60	Fairly difficult	Quality magazines	14 (14.1%)
61-70	Standard	Reader's digest	10 (10.1%)
71-80	Fairly easy	Slick fiction	3 (3.0%)
81-90	Easy	Pulp fiction	0 (0%)
91-100	Very easy	Comics	0 (0%)
* Flesch Reading Ease scores [12]			

The studied websites were then analyzed by the location and type of organization that developed the website. By location, the studied websites were 75.8% US, 7.1% European (UK), 6.1% Canadian, and 11% other (4% New Zealand, 3% Australia, 2% India, 1% South Africa, 1% Egypt). Analyzed for FRE scores by the four location groups, means of all websites fell into the 'difficult' category. By type, there were

51.5% commercial, 19.2% not-for-profit, 16.2% individual, 6.1% organization, 5.1% educational websites. Analyzed for FRE score by type, the average readability level of commercial websites was 40.7, with most commercial websites (60.8%) falling into the 'difficult' category, as observed for most not-for-profit (36.8%) and individual (56.3%) websites, and all (2) governmental sites; three organization websites (50%) were of the 'very difficult' level, and 2 (40%) of all educational websites were of the 'fairly difficult' level.

4. Discussion

Readability is a measure of the level of difficulty of a written text, and often predicts the extent to which a text is comprehensible to a target population. In any given population, there will be a mix of reading abilities [13]. For example, the reading level of the average US resident is grade 8 [14]. In the health care context, studies among publicly-insured individuals (e.g. Medicaid and Medicare recipients) demonstrate that the average reading level may be closer to grade 5. In the context of diabetes, studies have shown that the burden of diabetes tends to fall on the socioeconomically disadvantaged and the elderly, both groups that have average reading levels also lower than the US average [15]. From the public health perspective, health literacy may represent an important variable explaining the prevalence of poor health outcomes among patients with type 2 diabetes [16], as well as some of the socioeconomic, racial, and ethnic disparities in diabetes outcomes in the United States [17,18]. A considerable proportion of patients with type 2 diabetes is likely to have poor health literacy. In the United States, nearly 80% of patients with type 2 diabetes have completed only high school or less compared with 40% of the general population [19]. In this study, 75.8% websites were of US origin, so we used the average reading level of 8 when interpreting data.

Many studies have shown that most patient education literature written in English fails to conform to the current standards of readability [5], and the same holds true for information found on the Internet [6-10]. The discordance between readability and the average patients' capacities has been demonstrated across diverse patient populations, non-English languages [20], different health topics, and different forms of presentation or media used to provide health information [11].

Since HONcode accredited websites have been shown to have higher accuracy and reliability of Internet health information [2], they are often recommended to patients. We attempted to measure and additional aspect of quality, i.e. accessibility, by measuring the readability of HONcode accredited diabetes websites.

In this study, the readability levels for health information on 99 HONcode accredited diabetes websites ranged from 2.1 to 79.6 with the mean of 41.7, and median of 40.8. The obtained mean values are consistent with those found by other authors using FRE scores to test readability of health information on the Internet [6-10].

With the reading ability of the adult population set at 8, we can conservatively estimate that 86.9% of the studied websites provided information that would be too difficult for the average reader. The 'difficult' and 'very difficult' levels, i.e. of the academic or scientific levels, were found in 72.7% of the tested materials indicating either that the information was intended for physicians or that information from scientific writings was used for the websites by developers unaware of the need to adapt these materials to the reading ability of the intended readers.

Taking into account the frequent perception that the average quality of printed patient information leaflets is higher when written by pharmaceutical industry than by individual researchers [21], we assumed the same higher quality of Internet information presented by commercial websites. However, analysis by both the type and the location of website developers showed poor readability levels.

Although the reading levels of Web users tend to be somewhat higher than that of patients in general [10], with the growing number of Web users worldwide, adaptation of health information to the recommended reading levels becomes a necessity. Since readability formulas can serve as a guide to evaluate health information [14], producers should, at a minimum, test materials for readability and improve readability before posting on the Internet. While the goal or standard reading level often is described as the reading level of the "average reader", arguably the average reading level may vary in relationship to demographics of the target population. In addition, comprehensibility of text can be significantly improved by involving the target audience at all stages of material development.

It has been shown that both readability and understanding of text can be improved by increased attention to the linguistic features of the information [21]. Recently, we have witnessed a growing number of initiatives aimed at increasing awareness of written text producers about the literacy levels of the average adult ('Plain English Campaign' providing guides for writing medical information and tips for clear websites [22], 'Health Literacy Month' promoting understandable health information around the world [23], etc.). Even scientific journal editors, being aware of the inevitable change in the interaction between the public and the medical profession, wish to appeal to a wider audience through 'a more patient-friendly publication' [24], and some already include readability scores for information designed for patients [25].

Our minimum recommendation for assessment of quality of patient information on the Internet would be to indicate the readability level (which could be calculated by a sotfware tool thus making the process automated), and the name of the readability formula used on the HONcode accredited websites, as recommended for Internet sites in general [8], to serve as an indicator of the complexity of writing to all Web users searching the Internet for health information.

5. Conclusion

The readability levels obtained in this study ranged from FRE scores 2.1 to 79.6, with the mean 41.7 (10^{th} grade, 8^{th} month Flesch-Kincaid level), and indicate to a very complex writing not adapted to the reading ability of the general population. Providers of Internet health information designed for populations at large should be aware of the limited reading ability of their potential readers, and of the need to adapt their writing to the appropriate readability levels. It would be useful for consumers to have the readability level (and the name of the readability formula used) indicated on the website itself, along with the HONcode, to help them decide which sites could be of the greatest benefit for them.

6. References

[1] Eaton L. A third of Europeans and almost half of Americans use internet for health information. *BMJ* 2002;325:989

[2] Fallis D, Fricke M. Indicators of Accuracy of Consumer Health Information on the Internet. *JAMIA* 2002;9(1):73-79
[3] Health On the Net Foundation. Available at: http://www.hon.ch/HONcode. Accessed: October 31, 2002
[4] Zakaluk BL, Samuels SJ. eds. *Readability: Its Past, Present and Future*. Newark: International Reading Association, 1988
[5] Albert T. Written information for patients. *Lancet* 2000;356:434
[6] Graber MA, D'Alessandro DM, Johnson-West J. Reading Levels of Privacy Policies on Internet Health Web Sites. *J Fam Pract* 2002;51(7):642-645
[7] Smart JM, Burling D. Radiology and the Internet: a systematic review of patient information resources. *Clinical Radiology* 2001;56(11):867-870
[8] 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.
[9] 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. *Oncology Nursing Forum* 2000;27(9):1403-1409
[10] Graber MA, Roller CM, Kaeble B. Readability Levels of Patient Education Material on the World Wide Web. *J Fam Pract* 1999;48(1):58-61
[11] Ratzan SC, Parker RM. Introduction - Health literacy (bibliography online). Bethesda, MD: National Library of Medicine, 2000. Available at: http://www.nlm.nih.gov/pubs/resources.html
[12] Flesch R. A new readability yardstick. *J Appl Psychol* 1948;32:221-233
[13] Smith S. Readability Testing Health Information. Available at: http://www.prenataled.com. Accessed: October 20, 2001
[14] Mathews PJ et al. The Use of Readability Formulas in Patient Education Materials. Paper presented at the American Association for Respiratory Therapy Summer Forum, Reno, NV, July 26-28, 1985
[15] Schillinger D. Association of Health Literacy With Diabetes Outcomes. JAMA 2002;288:475-482
[16] Harris MI. Health care and health status and outcomes for patients with type 2 diabetes. *Diabetes Care* 2000;23:754-758
[17] Diabetes. In: *Healthy People 2010*. Washington, DC: Office of Disease Prevention and Health Promotion, US Dept of Health and Human Services, 2001
[18] Harris MI, Cowie CC, Reiber G, Boyko E, Stern M, Bennett P, eds. *Diabetes in America*, 2nd ed. Washington, DC: US Government Printing Office; 1995. Publication 95-1468
[19] Cowie CC, Eberhardt MS. Sociodemographic characteristics of persons with diabetes. In: Harris MI, Cowie CC, Reiber G, Boyko E, Stern M, Bennett P, eds. *Diabetes in America*, 2nd ed. Washington, DC: US Government Printing Office; 1995:85-116. Publication 95-1468
[20] Kusec S, Mastilica M, Pavlekovic G, Kovacic L. Readability of Patient Information on Diabetes on the Croatian Web Sites. In: G Surjan et al., eds. *Medical Informatics Europe 2002*. IOS Press 2002;128-132
[21] Bjorn E, Rossel P, Holm S. Can the written information to research subjects be improved? – an empirical study. *J Med Ethics*, 1999;25:263-267
[22] Plain English Campaign. Available at: www.plainenglish.co.uk. Accessed: October 30, 2002
[23] Health Literacy Month. Available at: www.healthliteracymonth.com. Accessed October 30, 2002
[24] Eaton L. A theme issue for medics and an increasingly health informed public. *BMJ* 2002;325:984
[25] Koops L, Lindley RI. Data Supplement - Patient Information Sheet. *BMJ* 2002;325(7361):415 Available at: www.bmj.com. Accessed: October 13, 2002

7. Address for correspondence

Sanja Kusec, Department of Educational Technology, Andrija Stampar School of Public Health, Medical School, University of Zagreb, Rockefellerova 4, HR-10000 Zagreb, Croatia; E-mail: sanja.kusec@snz.hr