American Journal of Men's Health
2018, Vol. 12(5) 1665–1669

DOI: 10.1177/1557988318780864
journals.sagepub.com/home/jmh
SAGE

Readability of Prostate Cancer Information Online: A Cross-Sectional Study

Corey H. Basch, EdD, MPH[1], Danna Ethan, EdD, MSW[2], Sarah A. MacLean, BA[3], Joseph Fera, PhD[4], Phillip Garcia[1], and Charles E. Basch, PhD[5]

Abstract

Reading and understanding health information, both components of health literacy, can influence patient decisions related to disease management. Older adults, the population of males at greatest risk for prostate cancer, may have compromised capacity to understand and use health information. The purpose of this study was to determine the readability of prostate cancer materials on the Internet using five recommended readability tests. Using a cleared Internet browser, a search was conducted for "prostate cancer." The URLs of the first 100 websites in English were recorded to create the sample. The readability scores for each website were determined using an online, recommended service. This service generates five commonly recommended readability tests. All five tests revealed that the majority of websites had difficult readability. There were no significant differences identified between websites with .org, .gov, or .edu extension versus those with .com, .net, or other extension. It is apparent that the Internet is used often as a resource for health-related information. This study demonstrates that the large majority of information available on the Internet about prostate cancer will not be readable for many individuals.

Keywords

prostate cancer, oncology/cancer, readability, online information

Received March 12, 2018; revised April 20, 2018; accepted May 8, 2018

In 2018, there will be an estimated 164,690 new cases of prostate cancer resulting in 29,430 deaths (American Cancer Society, n.d.). Prevalence is higher among males older than 65 years; African Americans are at particularly high risk for both prostate cancer incidence and mortality (Centers for Disease Control and Prevention [CDC], 2016a). In 2014, there were approximately 3,000,000 men living with prostate cancer in the United States (National Cancer Institute, n.d.). There has been considerable controversy about screening and treatment for prostate cancer (Carlsson & Vickers, 2015; Kim & Andriole, 2015; Tabayoyong & Abouassaly, 2015). It is therefore not surprising that the general public, at-risk males in particular, may be confused about whether they should participate in screening and if screening results are positive, make decisions about diagnostic tests and treatment options that warrant consideration. While these decisions should be informed primarily by individualized conversations between patients and their health-care provider(s), the Internet has become an increasingly popular media channel to which people turn to learn about

[1]Department of Public Health, William Paterson University, Wayne, NJ, USA
[2]Department of Health Sciences, Lehman College, The City University of New York, New York, NY, USA
[3]Icahn School of Medicine at Mount Sinai, New York, NY, USA
[4]Department of Mathematics, Lehman College, The City University of New York, New York, NY, USA
[5]Richard March Hoe Professor of Health and Education, Department of Health and Behavior Studies, Teachers College, Columbia University, New York, NY, USA

Corresponding Author:
Danna Ethan, Ed.D., MSW, Department of Health Sciences, Lehman College, CUNY, 250 Bedford Park Boulevard West, Gillet Hall, Rm. 431, Bronx, NY 10956, USA.
Email: danna.ethan@lehman.cuny.edu

health (Fox, 2014) and cancer (Foroughi, Lam, Lim, Saremi, & Ahmadvand, 2016).

Reading and understanding health information, both components of health literacy, can influence patient decisions related to disease management (Ratzan & Parker, 2000). The CDC defines health literacy as "the degree to which an individual has the capacity to obtain, communicate, process, and understand basic health information and services to make appropriate health decisions" (CDC, 2016b). One way to do this is to ensure that health materials are accessible to the intended audience. For all causes of death, there is a significant association between low health literacy and higher mortality (Fincham, 2013). A systematic review of roughly 100 articles confirmed that lower health literacy was related to higher mortality rates as well as other issues such as higher use of emergency care and greater rates of hospitalizations (Berkman, Sheridan, Donahue, Halpern, & Crotty, 2011). Specifically in men with newly diagnosed localized prostate cancer, one study found that those with lower health literacy had greater mental distress, suggesting that low health literacy also impacts quality of life (Song et al., 2012). While this may be explained, in part, by the colinearity between health literacy and levels of education and income, among other factors, access to comprehensible health information is necessary to help all individuals make informed decisions about their health (Roundtable on Health Literacy, 2014).

Older adults, the population of males at greatest risk for prostate cancer, may have compromised capacity to understand and use health information. Over 70% of persons aged older than 60 years have difficulty using print materials, according to the National Assessment of Adult Literacy, which was reported by the CDC (CDC, 2016c). By 2030, there will be an estimated 71.5 million adults aged 65 years and older in the United States, which highlights the importance of comprehensible health information about cancer (and other diseases and disorders) affecting this large and growing segment of the U.S. population (CDC, 2016c). It is recommended that health materials be written at or below the sixth-grade level (McKenzie, Neiger, & Thackeray, 2016). One study concluded that roughly half of those aged 55 to 64 years used the Internet to look up health information, whereby approximately one third of those aged 65 to 74 years reported using the Internet for this purpose. In this study, men and women aged 65 to 74 years reported virtually the same rates of Internet use for health purposes (Choi, 2011). Little is known about the readability of information on the Internet regarding prostate cancer. The purpose of this study was, therefore, to determine the readability of prostate cancer materials on the Internet using five recommended readability tests.

Methods

These methods were adapted from a previous study on the readability of online information (MacLean, Basch, Clark, & Basch, 2018). Using a cleared Internet browser (to clear the browser of cache, cookies, and history), a search was conducted for "prostate cancer." The browser used in this study was Google Chrome for speed, security, and simplicity. The URLs of the first 100 websites in English were recorded to create the sample. It should be noted that sponsored sites were not included in this sample.

The readability scores for each website were determined using Readable.io, which is a Medline-recommended service (National Institutes of Health, 2017). This service generates five commonly recommended readability tests: Flesch-Kincaid Grade Level (FKGL), Gunning Fog Index (GFI), Coleman-Liau Index (CLI), the Simple Measure of Gobbledygook (SMOG) Grade Level, and Flesch-Kincaid Reading Ease (FRE). The FKGL and FRE are calculated using sentence length and word length, but are weighted differently, with FRE deemed easier to read when scores are higher, and the FKGL scores align with grades of education in the United States (Readable.io, n.d.). Similarly, the results of the GFI test align with grades of education in the United States using a measure of average words per sentence and use of polysyllabic words (Readable.io, n.d.). CLI shares the same outcome of alignment with grades of formal U.S. education, but calculations depend on number of letters in a word and number of words in a sentence (Readable.io, n.d.). Like other tests, the SMOG Grade Level test uses a syllable counting system with a subset of text to determine the educational attainment needed to understand the material (McLaughlin, 1969). Based on the scores from these scales, websites were designated as having a readability level of "easy" (grade level <6), "average" (6–10), or difficult (>10).

In addition to descriptive statistics, it was determined whether or not there was an association between the readability of the information and the type of website. Websites were categorized as Group 1(.org, .gov, .edu) or Group 2 (.com, .net, other). SPSS (v23) and Microsoft Excel were used for data entry, data analysis, and to calculate descriptive statistics. Statistical tests included χ^2 tests of association for categorical variables (number of sites categorized as easy, average, or difficult on each for Group 1 vs. Group 2) and independent sample t-tests for continuous variables (mean score on each scale for Group 1 vs. Group 2). When an expected cell count was less than 5, Fisher's exact tests were used instead of χ^2 tests. Results were considered to be significant if $p < .05$. Descriptive statistics were used to depict the relationship between a site's general search position and its mean score on each readability test. The institutional review

Table 1. Descriptive Statistics on the Readability Tests of All Sites (n = 100).

Test	Minimum	Maximum	n			Mean	*SD*
			Easy: Grade <6	Average: Grade 6–10	Difficult: Grade >10		
FKGL	5.5	17.0	1	43	56	11.1	2.7
GFI	2.8	19.3	3	19	78	12.4	3.2
CLI	7.7	17.5	0	18	82	12.1	2.3
SMOG	8.0	17.8	0	11	89	12.7	2.3
FRE[a]	1.2	69.7	0	19	81	45.8	14.5

Note. FKGL = Flesch-Kincaid Grade Level; GFI = Gunning Fog Index; CLI = Coleman-Liau Index; SMOG = Simple Measure of Gobbledygook; FRE = Flesch-Kincaid Reading Ease.
[a]Scoring is as follows: Easy: score 80 to 100; average: score 60 to 79; difficult: score 0 to 59.

Table 2. Comparison of Websites Based on URL Type.

Test	Group 1[a] (n = 37)	Group 2[b] (n = 63)	p*	Group 1			Group 2			p**
				Easy	Avg.	Diff.	Easy	Avg.	Diff.	
FKGL	11.2	11.1	.941	0	17	20	1	26	36	.896
GFI	12.4	12.3	.964	1	9	27	2	10	51	.744
CLI	12.3	12.1	.670	0	4	33	0	14	49	.152[X]
SMOG	12.6	12.7	.895	0	6	31	0	5	58	.320
FRE	45.9	45.7	.940	0	8	29	0	11	52	.609[X]

Note. Avg. = average; FKGL = Flesch-Kincaid Grade Level; FRE = Flesch-Kincaid Reading Ease; SMOG = Simple Measure of Gobbledygook; GFI = Gunning Fog Index; CLI = Coleman-Liau Index; SMOG = Simple Measure of Gobbledygook; FRE = Flesch-Kincaid Reading Ease.
[a]Group 1: .org, .gov, .edu. [b] Group 2: .com, .net, other.
*Independent sample t-test. **Fisher's exact test.
[X]χ^2 test.

boards at William Paterson University and Teachers College, Columbia University deemed this study exempt.

Results

All five tests revealed that the readability of the majority of websites was difficult (Table 1). Based on the SMOG, 89% of websites were graded as difficult and the remaining 11% were graded as average. One of the readability tests indicated that 1 of the 100 websites had "easy" readability (FKGL), while another demonstrated that 3 websites were "easy" to read (GFI).other three readability tests did not find any of the 100 websites to be "easy" to read. Among the four tests that determine readability based on grade level, all found the average grade to be above the 11th grade, which indicates difficult readability.

There were no significant differences identified between Group 1 websites (.org, .gov, or .edu) and Group 2 websites (.com, .net, other; Table 2). This demonstrates that, regardless of URL type, most websites had difficult readability. Google Chrome often returns 10 search results per page. Thus, results were organized and analyzed by search position using a class width of 10. Table 3 shows the mean score of each search position group on each of the five readability tests. Each mean test score indicates difficult readability (FKGL, GFI, CLI, and SMOG score >10 or an FRE score <59). Since group 1 to 10 roughly corresponds to page 1 results, group 11 to 20 to page 2 results, and so forth, this table demonstrates that for any page, the expectation is a search result of difficult readability.

Discussion

The readability of prostate cancer treatment options on 62 websites revealed that there was a paucity of information written at below a high school reading level (Ellimoottil, Polcari, Kadlec, & Gupta, 2012). The findings of this study are consistent with that study and demonstrated that, more than 5 years later, the finding of inappropriate levels of readability remains consistent, even when using different readability assessment methods. While this study focused on readability of prostate cancer information, a study of patient information on the National Cancer Institute (NCI)-Designated Cancer Center websites was found to be, on average, at a twelfth grade

Table 3. Comparison of Mean Readability Score and Search Position Order.

	Mean readability score				
Search position	FKGL	GFS	CLI	SMOG	FRE
1–10	10.17	11.41	11.51	11.57	52.42
11–20	10.75	12.2	12.02	12.58	46.41
21–30	10.9	11.86	11.91	12.47	48.51
31–40	12.3	14.18	13.99	13.86	38.28
41–50	10.41	11.12	11.54	11.95	47.91
51–60	10.43	11.2	11.39	11.87	46.68
61–70	12.01	13.27	12.72	13.38	43.04
71–80	11.2	11.62	11.73	12.43	44.37
81–90	11.32	13.12	12.04	13.26	46.74
91–100	11.92	13.56	12.44	13.3	43.48

Note. FKGL = Flesch-Kincaid Grade Level; FRE = Flesch-Kincaid Reading Ease; SMOG = Simple Measure of Gobbledygook; GFI = Gunning Fog Index; CLI = Coleman-Liau Index.

reading level, despite the National Institutes of Health (NIH) recommendations that information be written at a sixth-grade level (Rosenberg et al., 2016). Collectively, these studies indicate that, in too many cases, the reading level required to access health information on the Internet may result in the information being inaccessible to the intended audience.

The Internet has become an increasingly important source of information related to cancer (Eysenbach, 2003). Patients diagnosed with cancer and their caregivers often seek out information about varied aspects of the disease (Basch, Thaler, Shi, Yakren, & Schrag, 2004). One study reported that 32% of prostate cancer patients utilized the Internet as a source of information gathering (Smith et al., 2003). Approximately 48% of cancer patients in one study felt that they had inadequate information about their cancer (Eysenbach, 2003). It is estimated that over 70% of adult Internet users search for health information online (Fox, 2013).

The limitations of this study include (a) the cross-sectional design, and (b) that reliance on the first 100 websites, an arbitrary cutoff point, on a search engine could yield different results at different points in time, given the Internet's constant flux. In addition, other tests, such as cohesion, were not performed and could provide further insight into understandability. Nonetheless, the findings from this study further affirm the need to create patient materials that are understandable for the general public about this prevalent, consequential, and controversial form of cancer.

Conclusion

It is apparent that the Internet is used often as a resource for health-related information. Given the controversial nature of prostate cancer screening and treatment options, it seems likely that consumers turn to the Internet for clarification information. Based on this study, the large majority of information available on the Internet about prostate cancer will not be readable for many individuals attempting to use the Internet to help inform their decision-making.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Corey H. Basch https://orcid.org/0000-0003-4862-4229

References

American Cancer Society. (n.d.). *Key statistics for prostate cancer. Prostate cancer facts*. Retrieved from https://www.cancer.org/cancer/prostate-cancer/about/key-statistics.html

Basch, E. M., Thaler, H. T., Shi, W., Yakren, S., & Schrag, D. (2004). Use of information resources by patients with cancer and their companions. *Cancer*, *100*(11), 2476–2483.

Berkman, N. D., Sheridan, S. L., Donahue, K. E., Halpern, D. J., & Crotty, K. (2011). Low health literacy and health outcomes: An updated systematic review. *Annals of Internal Medicine*, *155*, 97–107.

Carlsson, S., & Vickers, A. (2015). Spotlight on prostate cancer: The latest evidence and current controversies. *BMC Medicine*, *13*, 60.

Centers for Disease Control and Prevention. (2016a). *Basic information about prostate cancer*. Retrieved from https://www.cdc.gov/cancer/prostate/basic_info/index.htm

Centers for Disease Control and Prevention. (2016b). *What is heath literacy?* Retrieved from https://www.cdc.gov/healthliteracy/learn/index.html

Centers for Disease Control and Prevention. (2016c). *Importance of health literacy*. Retrieved from https://www.cdc.gov/healthliteracy/developmaterials/audiences/olderadults/importance.html

Choi, N. (2011). Relationship between health service use and health information technology use among older adults: Analysis of the US national health interview survey. *Journal of Medical Internet Research*, *13*(2), e33.

Ellimoottil, C., Polcari, A., Kadlec, A., & Gupta, G. (2012). Readability of websites containing information about prostate cancer treatment options. *Journal of Urology*, *188*(6), 2171–2175.

Eysenbach, G. (2003). The impact of the internet on cancer outcomes. *CA: A cancer journal for clinicians*, *53*(6), 356–371.

Fincham, J. E. (2013). The public health importance of improving health literacy. *American Journal of Pharmaceutical Education*, *77*(3), 41.

Foroughi, F., Lam, A. K., Lim, M. S., Saremi, N., & Ahmadvand, A. (2016). "Googling" for cancer: An infodemiological assessment of online search interests in Australia, Canada, New Zealand, the United Kingdom, and the United States. *JMIR Cancer*, *2*(1), e5.

Fox, S. (2014). *Pew Research Center: The social life of health information*. Retrieved from http://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/

Fox, S., & Duggan, M. (2013). *Pre research center: Health online*. Retrieved from http://www.pewinternet.org/2013/01/15/health-online-2013/

Kim, E. H., & Andriole, G. L. (2015). Prostate-specific antigen-based screening: Controversy and guidelines. *BMC Medicine*, *13*, 61.

MacLean S., Basch C.H., Clark A., & Basch C.E. (2018). Readability of information on colonoscopy preparation on the internet. *Health Promotion Perspectives*, *8*(2), 167–170.

McKenzie, J. F., Neiger, B. L., & Thackeray, R. (2016). *Planning, implementing & evaluating health promotion programs: A primer*. New York, NY: Pearson.

McLaughlin, G. (1969). SMOG grading: A new readability formula. *Journal of Reading*, *12*(8), 639–646.

National Institutes of Health. (2017). *How to write easy-to-read health materials*. Retrieved from http://www.nlm.nih.gov/medlineplus/etr.html

National Cancer Institute. (n.d.). *Prostate cancer: Cancer stat facts*. Retrieved from https://seer.cancer.gov/statfacts/html/prost.html

Ratzan, S. C., & Parker, R. M. (2000). Introduction. In C. R. Selden, M. Zorn, S. C. Ratzan, & R. M. Parker (Eds.), *In National Library of Medicine current bibliographies in medicine: Health literacy*. Bethesda, MD: National Institutes of Health. (NLM Pub. No. CBM 2000–1)

Readable.io. (n.d.). *Measure text readability – Readability and keyword density analysis and tools – Unique readability tools to improve your writing! Readable.io*. Retrieved from https://readable.io/

Rosenberg, S. A., Francis, D., Hullett, C. R., Morris, Z. S., Fisher, M. M., Brower, J. V., & … Kimple, R. J. (2016). Readability of online patient educational resources found on NCI-designated cancer center web sites. *Journal of the National Comprehensive Cancer Network*, *14*(6), 735–740.

Roundtable on Health Literacy. (2014). *Implications of health literacy for public health: Workshop summary*. Washington, DC: National Academies Press.

Smith, R. P., Devine, P., Jones, H., DeNittis, A., Whittington, R., & Metz, J. M. (2003). Internet use by patients with prostate cancer undergoing radiotherapy. *Urology*, *62*, 273–277.

Song, L., Mishel, M., Bensen, J. T., Chen, R. C., Knafl, G. J., Blackard, B., … Godley, P. A. (2012). How does health literacy affect quality of life among men with newly diagnosed clinically localized prostate cancer? Findings from the North Carolina-Louisiana Prostate Cancer Project (PCaP). *Cancer*, *118*(15), 3842–3851.

Tabayoyong, W., & Abouassaly, R. (2015). Prostate cancer screening and the associated controversy. *Surgical Clinics of North America*, *95*(5), 1023–1039.