



HHS Public Access

Author manuscript

J Natl Compr Canc Netw. Author manuscript; available in PMC 2020 May 19.

Published in final edited form as:

J Natl Compr Canc Netw. 2016 June ; 14(6): 735–740. doi:10.6004/jnccn.2016.0075.

Readability of Online Patient Educational Resources Found on NCI-Designated Cancer Center Websites

Stephen A. Rosenberg, MD¹, David Francis, MD¹, Craig R. Hullett, MD PhD¹, Zachary S. Morris, MD PhD¹, Michael M. Fisher¹, Jeffrey V. Brower, MD PhD¹, Kristin A. Bradley^{1,2}, Bethany M. Anderson, MD^{1,2}, Michael F. Bassetti, MD PhD^{1,2}, Randall J. Kimple, MD PhD^{1,2,*}

¹Department of Human Oncology, University of Wisconsin School of Medicine and Public Health, University of Wisconsin – Madison, Madison, WI, USA

²UW Carbone Comprehensive Cancer Center, University of Wisconsin School of Medicine and Public Health, University of Wisconsin – Madison, Madison, WI, USA

Abstract

Background: The NIH and Department of Health and Human Services recommend online patient information (OPI) be written at a 6th grade level. We used a panel of readability analyses to assess OPI from NCI Designated Cancer Center (NCIDCC) websites.

Methods: Cancer.gov was utilized to identify 68 NCIDCC websites from which we collected both general OPI and OPI specific to breast, prostate, lung, and colon cancer. This text was analyzed by 10 commonly used readability tests: the New Dale–Chall Test, Flesh Reading Ease Score, Flesh-Kinkaid Grade Level, FORCAST test, Fry Score, Simple Measure of Gobbledygook, Gunning Frequency of Gobbledygook, New Fog Count, Raygor Readability Estimate, and Coleman-Liau Index. We tested the hypothesis that the readability of NCIDCC OPI was written at the 6th grade level. Secondary analyses were performed to compare readability of OPI between comprehensive and non-comprehensive centers, by region, and to OPI produced by the American Cancer Society (ACS).

Results: A mean of 30,507 words from 40 comprehensive and 18 non-comprehensive NCIDCCs was analyzed (7 non-clinical and 3 without appropriate OPI were excluded). Using a composite grade level score, the mean readability score of 12.46 (i.e. college level, 95% CI: 12.13 – 12.79) was significantly greater than the target grade level of 6 (middle-school, $p < .001$). No difference between comprehensive and non-comprehensive centers was identified. Regional differences were identified in 4 of the 10 readability metrics ($p < .05$). ACS OPI provides easier language, 7–9th grade level, across all tests ($p < .01$).

Conclusions: OPI from NCIDCC websites is more complex than recommended for the average patient.

*CORRESPONDING AUTHOR: Randall J. Kimple, M.D., Ph. D., 3107 WIMR, 1111 Highland Ave., Madison, WI 53705, USA, Phone: (608) 265-9156, Fax: (608) 262-7224, rkimple@humonc.wisc.edu.

Publisher's Disclaimer: Disclaimers (by author): SR (None); DF (None); CH (None); ZM (None); MF (None); JB (None); KB (UptoDate-Intellectual Property Contributor); BA (Elekt); MB (None); RK (Threshold Pharmaceuticals)

Previously Presented: Presented at ASCO 2015, Chicago, IL. (ASCO Travel Award Recipient)

Introduction

Patient centered medicine is a cornerstone of 21st century healthcare in the United States. At the center of this ideal lies shared decision making between physician and patient. This process is essential in oncology, where patients are often confronted with an array of treatment options described by multiple specialists. These options are associated with individualized risks and benefits and are judged according to the unique values and interpretations of each patient. Patients often come to physician appointments with information gleaned from websites, the lay press, and online networking portals. There is clear evidence that patients are gathering treatment information on the Internet and are using that information to help guide their treatment decisions even before meeting with an oncologist¹. In order to make discerning oncologic treatment choices, it is imperative that patients find healthcare information from trusted and understandable sources. Poor patient comprehension of healthcare information correlates with lower patient satisfaction and compromises health outcomes²³. Patients place more trust in information and are more likely to follow recommendations that they understand²⁴. Effective communication plays an important role in overcoming health care disparities^{5,6}.

Comprehension of online healthcare information depends on a patient's health literacy; i.e. their ability to read and process health information and translate that information into healthcare decisions. Over one third of US adults have health literacy at or below the basic level⁷. Only 12% have proficient (i.e. the highest level) health literacy⁸. For comprehension, appropriately written healthcare information should account for average rates of health literacy in the US. National guidelines, including by the Department of Health and Human Services, recommend that health information be written at or below the 6th grade level based on the current US literacy rate⁹.

We set out to determine whether patient materials found on the web sites of NCI-Designated Cancer Centers (NCIDCC) were written at an appropriate level to facilitate patient comprehension. Other groups have demonstrated significant gaps between recommended information complexity and actual written information for patients in a variety of non-oncology fields¹⁰⁻¹⁴. We gathered and examined patient-targeted information found on NCIDCC websites and analyzed it with 10 distinct tests of readability. We tested the hypothesis that the average readability of OPI from NCIDCC websites would be written at the appropriate 6th grade level. We performed secondary exploratory analyses to investigate potential differences in readability between geographic regions or between comprehensive and non-comprehensive cancer centers. Results from freely available American Cancer Society OPI were used as a comparison. If OPI from NCIDCC were not written at an appropriate level for optimal patient comprehension, this finding would have significant implications for shared decision-making and healthcare disparities.

Methods:

Text extraction

We identified NCIDCC websites using the online list available at <http://cancercenters.cancer.gov/Center/CancerCenters> as of October 2014. Websites were

individually viewed by one of two study authors (DF, MF) and patient-targeted information related to general information, treatment options, and side effects for breast, prostate, lung, and colon cancer was extracted. Our analysis included information about general descriptions of each cancer, screening, treatment options (e.g. surgery, chemotherapy, radiation), benefits, side effects, risks, survivorship, and other issues surrounding cancer care. Links to scientific protocols, citations, references, patient accounts, physician profiles, or outside institutions were explicitly excluded from this analysis. Every attempt was made to capture online information from NCIDCC websites, as patients would have encountered it as reading through each appropriate website. In addition, any links leading outside each individual cancer center's domain (e.g. to the National Cancer Institute or American Cancer Society) were also excluded from analysis.

For comparison, text from OPI relevant to breast, prostate, lung and colon cancer available on the American Cancer Society (ACS) website as of October 2014 was also collected. The ACS OPI source material includes information that is often encountered through each NCIDCC website regarding: description of cancer, screening, diagnosis, treatment options (chemotherapy, surgery, radiation), and side effects. These documents are comprehensive and thorough in their content of the aforementioned topics. Because these documents are similar in the type and scope of information presented on NCIDCC websites, they were chosen as a choice of comparison.

Assessment of readability

Extracted text was uploaded into Readability Studio ® ver. 2012.0 for analysis (Oleander Software, Hadapsar, India). We chose 10 commonly used readability tests to assess the readability of this material and avoid potential biases present within each individual test: the New Dale-Chall Test, Flesh Reading Ease Score, Flesh-Kinkaid Grade Level, FORCAST test, Fry Score, Simple Measure of Gobbledygook (SMOG), Gunning Frequency of Gobbledygook, New Fog Count, Raygor Readability Estimate, and Coleman-Liau Index. These tests are used in both the public and private sector and have been well validated as measures of readability^{15–22}. Each test reports a score or score range which was utilized for all analyses.

Statistical analysis

Statistical analyses were performed using IBM SPSS for Macintosh, version 22.0 (IBM Corporation, Armonk, New York). We compared and measured readability of OPI from NCI Designated Cancer Centers to the 6th grade level. We chose a 6th grade reading level as the standard against which to compare the readability of the texts because this has been established as the target grade level by the Department of Health and Human Services⁹. Given a set standard for comparison, single-sample t-tests were used to determine the significance of the difference between our texts and the ideal reading level. Additional analyses included a comparison of comprehensive vs. non-comprehensive NCIDCC using independent-samples t-tests. We also assessed whether readability varied systematically according to geographic region, using ANOVA to determine significant differences. We used the geographic regions as defined by the National Adult Literacy survey²³. This divided

cancer centers into one of four census definitions of regions: Northeast, Midwest, South, and West. States for each region are listed in Supplemental Table 1.

Results:

OPI was collected from 58 NCIDCCs. Only non-clinical centers (n=7) and those without online patient information (n=3) were excluded (Figure 1). A mean of 30,507 words (range 5,732 – 147,425), 1,639 sentences (range 319–7094), and 762 paragraphs (range 155–2652) per website were extracted.

Two of the most commonly reported readability tests are the Flesch-Reading Ease scale and the Raygor Readability estimate. The Flesch-Reading Ease scale generates a score ranging from 100 (very easy) to 0 (very difficult) with “plain English” scoring a level of 60–70 (understood by most 13–15 year olds). This test focuses on words per sentence and syllables per word and is a standard measurement of readability often used by US government agencies¹⁸. The mean score on this test was 43.33 (sd 7.46, range 27–57) for comprehensive and 44.78 (sd 6.63, range 31–55) for non-comprehensive NCIDCC (Supplemental Figure 1A, B). In these analyses, OPI at all NCIDCC is at least two standard deviations away from the target goal for an appropriate reading level based on the Flesch-Reading Ease scale. The Raygor Readability estimate uses the number of sentences and letters per 100 words and provides a grade level estimate. Using this test, OPI for all NCIDCCs was 14.1 (i.e. college-level, sd 2.3, Figure 2). Again, this is significantly higher than the target goal of a 6th grade level ($p < .001$).

To bolster our analysis and because there is not one single validated health-care specific readability test, we analyzed OPI using a panel of 10 different tests. Across all 10 readability tests, OPI found at NCIDCC (whether comprehensive or non-comprehensive) was written at a significantly higher level than the target 6th grade level ($p < .05$). Eight scales (Coleman-Liau, Flesch-Kincaid, FORCAST, Fry, Gunning Fog, New Fog, Raygor, and SMOG) provided a single measure of grade level as an output (Table 1). These measures are highly inter-correlated, ranging from $r = 0.74$ to 0.98 , and in combination form a highly reliable scale (SI alpha = 0.99). We therefore used the average of these eight measures as a composite grade level reflecting readability. Across 58 centers, the mean of the eight tests reported a grade-level score of 12.46 (college level, 95% CI: 12.13 – 12.79), which was significantly higher than the target grade level of 6 (middle school) recommended by the Department of Health and Human Services, $t(57) = 38.15$, $p < .001$.

An additional set of analyses was performed to determine whether the reading levels of OPI from NCIDCCs were significantly different from those obtained from the ACS patient handouts. Across all metrics, ACS websites provide easier language (Table 1 and Figure 2, $p < .01$ in all cases). Post-hoc comparisons of the subgroup means using Bonferroni corrections indicated that the ACS mean was significantly different from both comprehensive and non-comprehensive cancer centers for each readability metric ($p < .05$ for each).

Across individual readability tests, there were no differences in readability between comprehensive and non-comprehensive cancer centers. (Table 1 and Supplemental Figure 1A, 1B). Similarly, no difference between the mean readability for comprehensive (12.56, sd 1.33) versus non-comprehensive cancer centers (12.24, sd 1.19) was found using the composite scale of eight measures of readability, $t(56) = 0.90$, $p = 0.37$.

Finally, as there are documented differences in literacy across geographic regions, we assessed regional differences in readability as an exploratory analysis. When comparing individual readability measures, significant regional differences were identified in 4 of the 10 metrics (Supplemental Table 2). Websites from the Midwest tended to provide information that was easiest to read, while those from the Northeast and West were the most difficult (Figure 3). Using the previously described composite measure of readability, only a trend toward regional variation was seen ($p=0.08$).

Discussion:

The past twenty years have witnessed a paradigm shift in how patients obtain health care information. What was once the domain of physicians and other health care providers is now often supplemented or replaced by keyword or symptom searches on the Internet¹. Patients searching for online health information can find helpful, balanced, and appropriate counseling but may also encounter misinformation, scare tactics, and highly biased reports. Patients seek out information they can comprehend and will often consider that information as a trusted source^{24,25}.

NCI-Designated Cancer Centers are in a position to be at the forefront of providing patients with access to appropriate cancer related information. In this analysis we determined how well patient information is presented based on recommended target readability levels. National guidelines recommend that health information be presented at a 6th grade reading level. We found that OPI from NCI-Designated Cancer Centers was written at nearly double that level (Table 1 and Figures 2). This corresponds to the reading level of a first year college student. A significant disconnect exists between what patients can understand and what is provided online by NCIDCC websites.

We focused our analyses on general patient information and educational material related to four of the most common cancers in the US: breast, colon, prostate and lung. We chose to use a panel of readability analyses rather than rely on a single measure of readability. These analyses have been well validated in a number of different settings. Regardless of test used, information from NCIDCCs was too complex based on the target level.

In discussing our preliminary findings with colleagues, it commonly was suggested that cancer care information is too complex to be written at a 6th grade level. Analysis of OPI provided by the ACS documents that information is written closer to the target grade level. ACS documents included information about chemotherapy, radiotherapy, and surgery and are written between a 7th and 9th grade level. Although still written above the 6th grade level, this information is presented at a more appropriate level than that seen at NCIDCC websites.

We performed a number of secondary analyses with our data. We compared readability of OPI between comprehensive and non-comprehensive cancer centers and found no statistically significant difference between these two groups. The designation of a comprehensive cancer center is determined by the NCI and is dependent on the services provided by the center to patients and on ongoing research. The quality of OPI is not considered for NCI designation. The National Adult Literacy Survey (NALS) has documented regional differences related to educational level, immigration, poverty and other factors²³. We placed cancer centers into regions as defined by the NALS and found regional differences in 4 of 10 tests. On average the grade level of OPI is highest in the Northeast (12.9 grade level) and lowest in the Midwest (11.8) (Figure 3). This contrasts with the results of the NALS where adults in the Midwest outperformed those in the Northeast in regards to literacy²³. This suggests that a larger gap between readability of information and literacy is seen in the Northeast than in other regions.

The majority of readability measures we used derived their metric (i.e. grade output) by analyzing the number of syllables per word and the number of words per sentence. This implies that the use of simpler words and/or shorter sentences would decrease the reading level of OPI. For example, “immunologic modulation via pharmacologic targeting of the PD1 receptor” could instead be written as “drugs that help your own immune system fight cancer”. This changes the grade level from doctorate to approximately 7th-9th grade.

Improving the information found on NCI Cancer Center websites requires a multidisciplinary approach involving physicians, nurses, healthcare educators, and patients. Experts in healthcare communication also should be included when developing OPI. An open dialogue and exchange focused on improving accuracy and access to this information and communication has the potential to improve outcomes and benefit patients²⁶. With this in mind, there must be a push forward in how to make improvements in how information is presented and disseminated to patients²⁷. Besides improving the actual content of information, appropriate separation of information for patients, clinicians, and researchers, could help greatly improve online communication and readability. Few NCIDCC websites have now created specific “sub-sections” to their website explicitly targeted to one the abovementioned groups. This could help ensure that complex clinical or scientific information is not incorporated into OPI that is meant for the lay public. We therefore advocate that this an important first step for each NCIDCC website to make in order to improve communication of online information for patients, clinicians, and researchers.

We acknowledge several limitations of this work. We only analyzed data from a single time point while cancer center websites can change daily, weekly, or monthly. While information on research findings or clinical trials may be updated regularly, it is less likely that educational material provided to patients is revised frequently. This is partly compensated by the analysis of data from every NCI-Designated Cancer Center. In addition, we did not attempt to assess the accuracy of content, only the readability of information. Viewing the websites it becomes clear that many websites mix patient information with information for clinicians and researchers. A number of centers have begun to provide information specific for each group of potential users. Such an approach should allow data to be presented at a reading level appropriate for their target audience. Finally, we did not attempt to analyze

printed patient educational materials that are often provided to patients visiting these centers. It is possible that this information is written at a more appropriate level.

NCI-Designated Cancer Centers are identified as the local and national centers of excellence where patient care, research, and innovation take place. Many patients look to their local NCI-Designated Cancer Center as a source of both information and treatment. It is imperative that patient information from these centers be well written, accurate, and understandable for the majority of Americans. Failure to provide appropriate information can result in patients seeking alternative sources of information, often of variable quality. NCI-Designated Cancer Centers should be leaders in disseminating accurate and appropriate written information on cancer care.

Conclusions

In this study, we examined the readability of OPI found on NCI-Designated Cancer Center websites. We found that information is written at the 12–13th grade level (freshmen collegiate level), which is significantly above the recommended national guidelines of 6th grade. Improvement in the readability of OPI will depend on a multidisciplinary involvement of physicians, nurses, educators, and patients.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Research Support:

This work was supported in part by NIH/NCI CA160639 (RJK).

References

1. Koch-Weser S, Bradshaw YS, Gualtieri L, et al.: The Internet as a health information source: findings from the 2007 Health Information National Trends Survey and implications for health communication. *J Health Commun* 15 Suppl 3:279–93, 2010 [PubMed: 21154099]
2. Bains SS, Bains SN: Health literacy influences self-management behavior in asthma. *Chest* 142:1687; author reply 1687–8, 2012 [PubMed: 23208348]
3. Health literacy: report of the Council on Scientific Affairs. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. *JAMA* 281:552–7, 1999 [PubMed: 10022112]
4. Rosas-Salazar C, Apter AJ, Canino G, et al.: Health literacy and asthma. *J Allergy Clin Immunol* 129:935–42, 2012 [PubMed: 22326486]
5. Berkman ND, Sheridan SL, Donahue KE, et al.: Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med* 155:97–107, 2011 [PubMed: 21768583]
6. Hirschberg I, Seidel G, Strech D, et al.: Evidence-based health information from the users' perspective--a qualitative analysis. *BMC Health Serv Res* 13:405, 2013 [PubMed: 24112403]
7. S. W: Assessing the nation's health literacy: Key concepts and findings of the National Assessment of Adult Literacy (NAAL) American Medical Association Foundation 2008
8. Cutilli CC, Bennett IM: Understanding the health literacy of America: results of the National Assessment of Adult Literacy. *Orthop Nurs* 28:27–32; quiz 33–4, 2009 [PubMed: 19190475]
9. NIH: How to Write Easy-to-Read Health Materials. National Institute of Health, 2013 <https://www.nlm.nih.gov/medlineplus/etr.html>

10. Colaco M, Svider PF, Agarwal N, et al.: Readability assessment of online urology patient education materials. *J Urol* 189:1048–52, 2013 [PubMed: 23017508]
11. Svider PF, Agarwal N, Choudhry OJ, et al.: Readability assessment of online patient education materials from academic otolaryngology-head and neck surgery departments. *Am J Otolaryngol* 34:31–5, 2013 [PubMed: 22959363]
12. Shukla P, Sanghvi SP, Lelkes VM, et al.: Readability assessment of internet-based patient education materials related to uterine artery embolization. *J Vasc Interv Radiol* 24:469–74, 2013 [PubMed: 23452554]
13. Misra P, Kasabwala K, Agarwal N, et al.: Readability analysis of internet-based patient information regarding skull base tumors. *J Neurooncol* 109:573–80, 2012 [PubMed: 22810759]
14. Eloy JA, Li S, Kasabwala K, et al.: Readability assessment of patient education materials on major otolaryngology association websites. *Otolaryngol Head Neck Surg* 147:848–54, 2012 [PubMed: 22864405]
15. Walsh TM, Volsko TA: Readability assessment of internet-based consumer health information. *Respir Care* 53:1310–5, 2008 [PubMed: 18811992]
16. Albright J, de Guzman C, Acebo P, et al.: Readability of patient education materials: implications for clinical practice. *Appl Nurs Res* 9:139–43, 1996 [PubMed: 8771859]
17. Cooley ME, Moriarty H, Berger MS, et al.: Patient literacy and the readability of written cancer educational materials. *Oncol Nurs Forum* 22:1345–51, 1995 [PubMed: 8539175]
18. FLESCHE R: A new readability yardstick. *J Appl Psychol* 32:221–33, 1948 [PubMed: 18867058]
19. GH. M: SMOG grading: a new readability formula. *J reading*, 1969
20. Coleman MLT: A computer readability formula designed for machine scoring. *J Appl Psychol*, 1975
21. AL. R: The Raygor Readability Estimate: a quick and easy way to determine difficulty, in In & Pearson PD e (ed): *National Reading Conference. Reading: theory rap.*, 1977, pp 259–263.
22. Fry E: A readability formula that saves time. *Journal of Reading* 11, 1969
23. Kirsch I, Jungeblut A, Jenkins L, et al.: *Adult Literacy in America*, (ed 3rd Edition), National Center for Education Statistics, 2002
24. Baker DW: The meaning and the measure of health literacy. *J Gen Intern Med* 21:878–83, 2006 [PubMed: 16881951]
25. Ellimoottil C, Polcari A, Kadlec A, et al.: Readability of websites containing information about prostate cancer treatment options. *J Urol* 188:2171–5, 2012 [PubMed: 23083852]
26. Safeer RS, Keenan J: Health literacy: the gap between physicians and patients. *Am Fam Physician* 72:463–8, 2005 [PubMed: 16100861]
27. Weiss BD: Health Literacy Research: Isn't There Something Better We Could Be Doing? *Health Commun* 30:1173–5, 2015 [PubMed: 26372029]

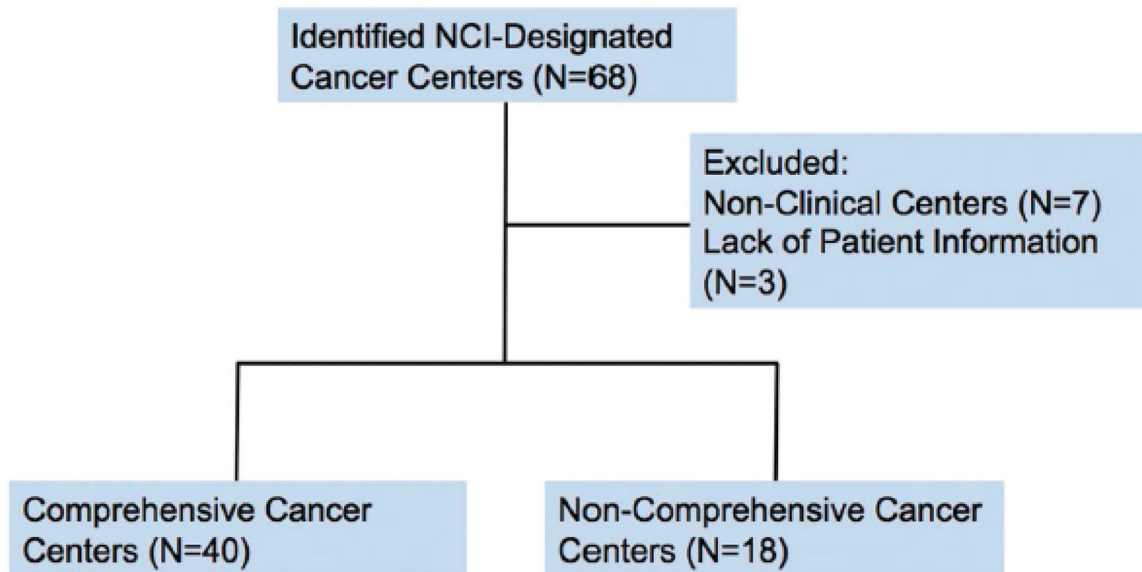


Figure 1: CONSORT diagram of study design and flow.

68 NCI-Designated Cancer Centers were identified. Non-clinical centers (n=7) and centers that lacked patient information (n=3) were excluded from the analysis.

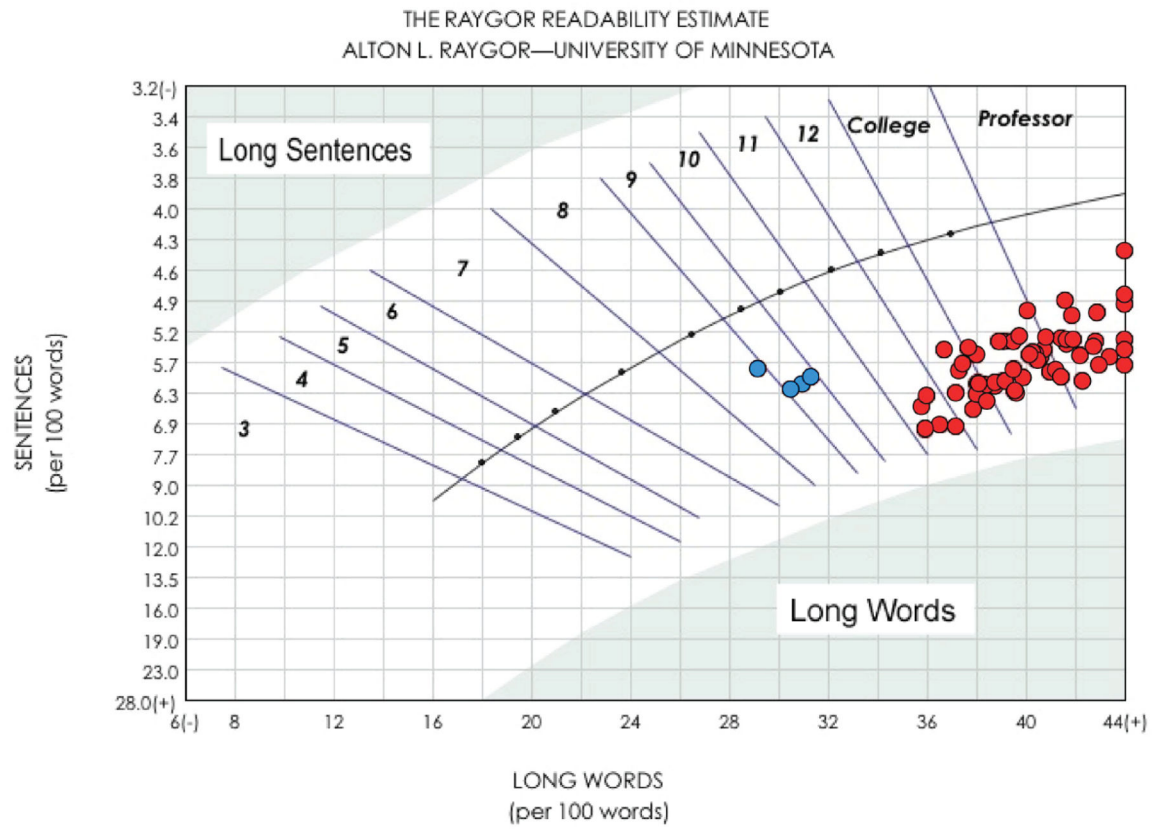


Figure 2: Raygor Readability level. OPI from NCIDCC websites (red) and ACS websites (blue) underwent Raygor Readability analysis. OPI for all NCIDCC was 14.1 (college-level, sd 2.3). The ACS OPI provides easier language (7–9th grade) compared to NCIDCC websites ($p < 0.01$).

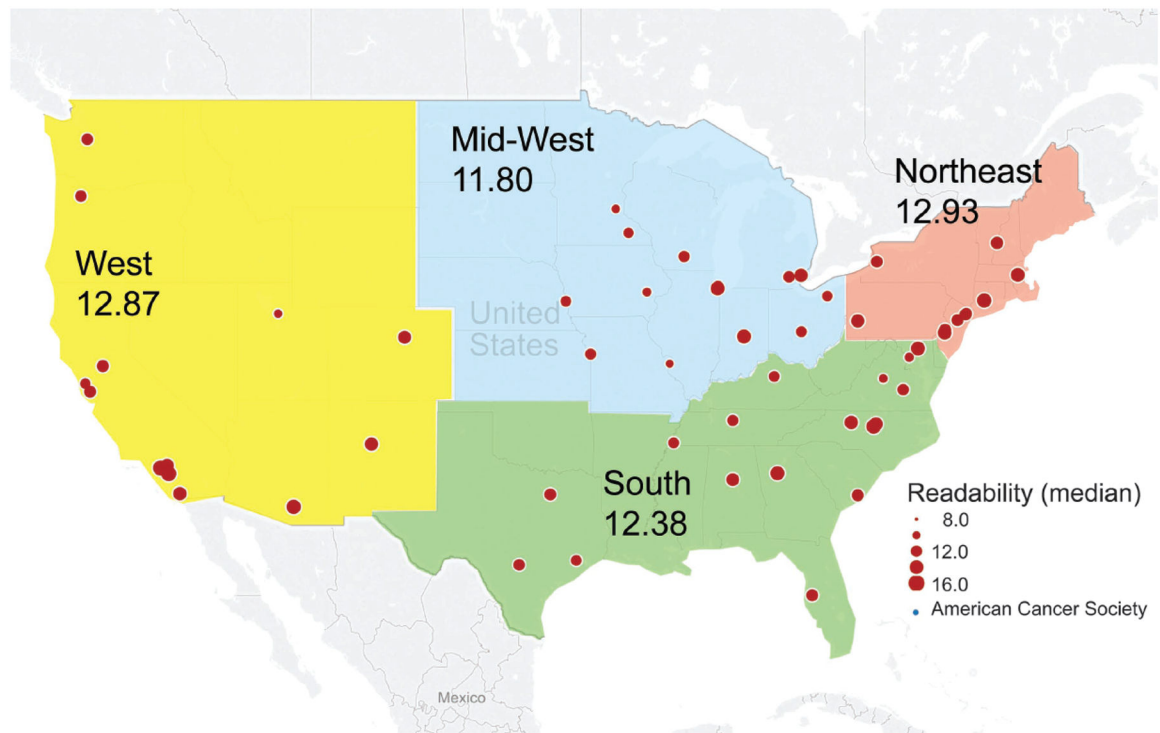


Figure 3: Regional variations in readability of OPI.

Cancer centers were assigned to regions based on the National Adult Literacy Survey (West=yellow, Mid-West=blue, South=green, and Northeast=orange). Individual NCIDCC are shown with circle size representing the composite grade level measure (red). For comparison, results of the ACS analysis are provided (legend, blue circle). Regional differences were identified in 4 of the 10 readability metrics ($p < 0.05$) with higher levels in the Northeast (12.9) and lower levels in the Mid-west (11.8)

Table 1.

Readability of NCIDCC and ACS Online Patient Information

| | Comprehensive (mean, SD) | Non-Comprehensive (mean, SD) | ACS (mean, SD) | p-value ^j (NCIDCC vs ACS) |
|----------------|--------------------------|------------------------------|----------------|--------------------------------------|
| Coleman-Liau | 12.84 (1.16) | 12.60 (1.10) | 9.35 (0.25) | <.001 |
| Flesch-Kincaid | 11.85 (1.52) | 11.39 (1.26) | 8.40 (0.18) | <.001 |
| FORCAST | 11.50 (0.40) | 11.46 (0.39) | 10.18 (0.15) | <.001 |
| Fry | 14.25 (2.15) | 13.94 (1.98) | 8.50 (0.58) | <.001 |
| Gunning Fog | 12.88 (1.29) | 12.64 (1.06) | 9.10 (0.14) | <.001 |
| New Fog | 9.29 (1.26) | 8.89 (1.07) | 7.32 (0.29) | 0.007 |
| Raygor | 14.28 (2.32) | 13.67 (2.25) | 8.75 (0.50) | <.001 |
| SMOG | 13.62 (1.18) | 13.29 (0.97) | 10.08 (0.21) | <.001 |

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript