




Investigation of the Readability and Reliability of Online Health Information for Cancer Patients During the Coronavirus Pandemic

Cameron Schluter^{1,2} · Maia Fefer² · Grace Lee^{3,4} · Isaac G. Alty^{3,5} · Edward Christopher Dee^{3,6} 

Accepted: 17 January 2022 / Published online: 25 April 2022
© The Author(s) under exclusive licence to American Association for Cancer Education 2022

Abstract

For cancer patients undergoing treatment who may be at higher risk of COVID-19, access to high-quality online health information (OHI) may be of particular importance amidst a plethora of harmful medical misinformation online. Therefore, we assessed the readability and quality of OHI available for various cancer types and treatment modalities. Search phrases included “cancer radiation COVID,” “cancer surgery COVID,” “cancer chemotherapy COVID,” and “cancer type COVID,” for the fourteen most common cancer types (e.g., “prostate cancer COVID” and “breast cancer COVID”), yielding a total of 17 search phrases. The first 20 sources were recorded and analyzed for each keyword, yielding a total of 340 unique sources. For each of these sources, the approximate grade level required to comprehend the text was calculated as a mean of five validated readability scores; subsequently, for the first ten results of each search, the DISCERN tool was manually used to assess quality. Search terms were translated into Spanish and French, and a quality assessment using the Health on the Net Code (HONcode) accreditation was conducted. The median grade level readability for all sources was 13 (IQR 11–14). Median DISCERN scores for the 170 sources assessed were 55 out of 75, suggesting good quality. OHI with quality scores below the median DISCERN score had a median readability of 12.5 (IQR 11–14) grade reading level vs 14 (IQR 12–17) for those above the median DISCERN score (T -test $P < 0.0001$). Percentages of HONcode-accredited websites were 34.9%, 39.9%, and 38.6% for English, Spanish, and French OHI, respectively. We conclude that efforts are needed to make high-quality OHI available at the appropriate reading level for patients with cancer; such efforts may contribute to the alleviation of disparities in access to healthcare information.

Keywords COVID-19 · Online health resources · Oncology · Patient education · Cancer education

Introduction

Perhaps accelerated by the COVID-19 pandemic where the need for physical distancing has reduced in-person patient-physician encounters, online health information (OHI) remains a critical contributor to patient education [1]. For cancer patients undergoing treatment who may be at higher risk of COVID-19 infection and symptoms of greater severity, access to high-quality OHI may be of particular importance [1, 2]. With the rising reliance of OHI for cancer patients in recent years [3], there is a growing necessity for reliable and accessible OHI in the midst of potentially harmful medical misinformation [2], particularly for cancer patients for whom early detection of tumors is important in providing treatment.

Readability—the approximate level of education needed to comprehend a piece of text—is an important aspect of accessible OHI [4–6]. The American Medical Association

✉ Isaac G. Alty
Isaac_Alty@BWH.Harvard.edu

✉ Edward Christopher Dee
DeeE1@mskcc.org

¹ Department of Biochemistry, Colby College, Waterville, ME, USA

² University of Massachusetts Medical School, Worcester, MA, USA

³ Harvard Medical School, Boston, MA, USA

⁴ Department of Radiation Oncology, Dana Farber Cancer Institute/Brigham and Women’s Hospital, Boston, MA, USA

⁵ Department of Surgery, Brigham and Women’s Hospital, Boston, MA, USA

⁶ Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

(AMA) suggests the provision of OHI at a sixth-grade reading level to appropriately serve the average US reader and to improve retention of health information [7]. As OHI has been shown to influence patient decisions [8], the importance of accessible and high-quality OHI is crucial, particularly during an era with an increasing dependence on virtual visits and online resources. We therefore aim to analyze the readability and quality of OHI available for cancer patients during the COVID-19 pandemic.

Given the harmful potential of misinformation surrounding the COVID-19 pandemic and the importance of proper treatment for cancer patients, it is important to understand and analyze the accessibility, reliability, and quality of OHI available for various cancer types and treatments as these diseases relate to COVID-19. We hypothesize that OHI regarding coronavirus for cancer patients would have highly variable quality and would require a grade reading level greater than the AMA recommended sixth-grade, possibly impacting patient decisions and comprehension regarding their treatment options.

Methods

Google Trends (Google, Mountain View, CA) was used to assess the United States search frequency of “coronavirus” and “COVID,” with the more commonly used term, “COVID,” selected. A focus group decided the three cancer treatment modalities and fourteen most common cancer types would constitute the search phrases, constructed as “cancer type/treatment modality + COVID,” yielding 17 search phrases (Table 1). Since patients are unlikely to search past the first two pages [9], the first 20 sources were recorded and analyzed for each search phrase, yielding 340 unique sources compiled on June 30, 2020.

These sources were categorized by a single author (C.G.S.) based on their classification as “news,” “medical sources for provider or research,” “medical sources for patient,” and “other.” The readability of each source was assessed with the readability test tool (WebFX, Harrisburg, PA) by examining the full text and calculating the reading level based on school grade. The WebFX instrument generates a composite readability grade level (CGL) that represents the rounded average of five validated readability formulas, namely the Flesch Kincaid Grade Level, the Gunning Fog Score, the SMOG Index, the Coleman Liau Index, and the Automated Readability Index [4, 10]. We chose to report the average of these five validated metrics (i.e., the CGL) to account for some variability in how these scores are calculated.

Quality was assessed for the first ten results of each search ($n = 170$) using the DISCERN instrument [6]. DISCERN consists of 15 questions scored out of 5 for a maximum

Table 1 Displays the observed median and interquartile range for grade reading level and DISCERN scores out of 75 for each English search phrase used. The AMA recommends a grade readability of 6 and the median DISCERN score that determined high- or low-quality sources was a score of 55

	Median readability (IQR)	Median DISCERN (IQR)
Cancer surgery COVID	13 (11,14)	54 (51,57)
Cancer chemotherapy COVID	13 (11,14)	54 (52,62)
Cancer radiation COVID	13.5 (12,15)	58 (54,64)
Bladder cancer COVID	13 (12,15)	55 (53,56)
Breast cancer COVID	12 (11,15)	58 (55,60)
Colon cancer COVID	14 (12.75,17)	57 (54,61)
Endometrial cancer COVID	14 (12,15.25)	57 (51,58)
Kidney cancer COVID	13 (12,16)	57 (54,63)
Leukemia COVID	14.5 (12,15.25)	62 (61,64)
Liver cancer COVID	14 (12.75,16)	56 (41,60)
Lung cancer COVID	12.5 (10.75,14)	57 (55,61)
Melanoma COVID	13 (11,15.25)	55 (47,55)
Non-Hodgkin lymphoma COVID	12.5 (11,14.5)	54 (50,57)
Pancreatic cancer COVID	14.5 (12,15.25)	55 (48,57)
Prostate cancer COVID	13.5 (12,15.5)	51 (50,53)
Rectal cancer COVID	16.5 (13.75,17)	53 (52,57)
Thyroid cancer COVID	12.5 (11,14)	52 (46,60)
News sources	13 (11,15)	54 (49,62)
Medical sources for provider	13 (11,15)	57 (54,61)
Medical sources for patient	13 (12,15)	55 (49,60)

score of 75 and is utilized widely to assess the reliability and quality of OHI for treatment options [10–12]. The two primary sections of the DISCERN survey are “is the publication reliable” (Section 1) and “how good is the quality of information on treatment choices” (Section 2). Two authors (C.G.S. and E.C.D.) independently scored each source, conferred, and agreed upon each DISCERN rating, while omitting the subjective final question of the DISCERN instrument “rate the overall quality of the publication.” DISCERN scores were then binarized at the observed median to label the sources collected as “high quality” or “low quality” and were then compared in relation to the sources’ grade level readability.

Lastly, a multilingual quality assessment, with each search phrase translated into Spanish and French, was conducted using the language-agnostic Health on the Net Code certification for the first 10 sources of each keyword to rapidly address and identify credible and reliable OHI [6, 13, 14]. The Health on the Net Foundation is a non-profit, non-governmental organization [6] formed with the intent of identifying and certifying reliable and credible OHI for

safe use on the internet. *t* tests, one-way analysis of variance (ANOVA), and Pearson correlation were used to compare readability and quality metrics with significance defined at $\alpha = 0.05$.

Results

The 340 sources collected and classified were composed of 19 news sources, 114 medical research sources, 164 medical resources for patient, and 43 other sources.

Median grade level readability for all sources was 13 (IQR 11–14). In terms of subgroups, the median grade level readability was 13 (IQR 11–15) for news sources, 14 (IQR 11–15) for medical research sources, 13 (IQR 12–15) for medical resources for patient, and 12 (11–14) for other sources. The readability scores of binarily

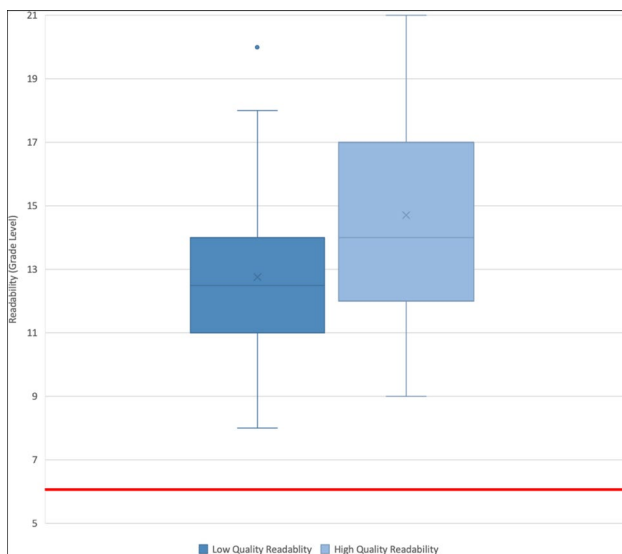
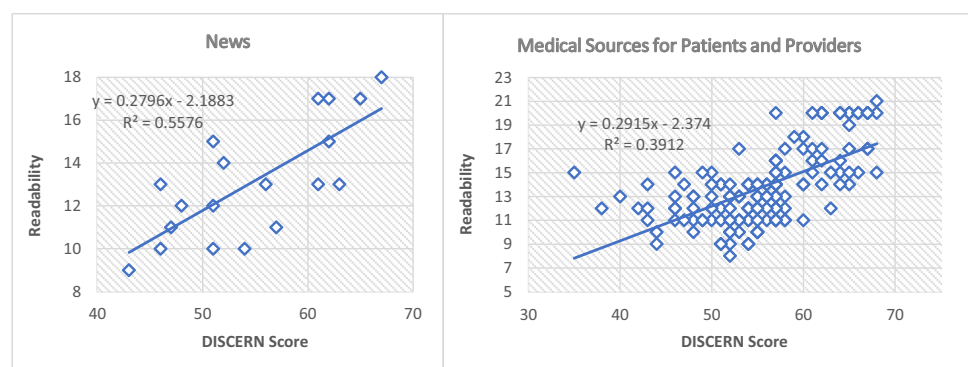


Fig. 1 Boxplot describing readability grade level of OHI stratified by quality binarized at median DISCERN score. Median and mean are marked within the boxplots. The red line indicates the desired sixth-grade reading level as determined by the AMA

Fig. 2 Scatterplot of readability versus quality of OHI for cancer and COVID for a line of best fit using news-related articles (a) and medical-related resources (b)



defined high- and low-quality sources demonstrated a statistically significant difference; low-quality OHI recorded a median 12.5 (IQR 11–14) grade reading level vs 14 (IQR 12–17) for high-quality OHI (*t*-test $p < 0.0001$, Fig. 1). Further analysis through Pearson correlation tests of the different source classes also showed a significant association between the quality and readability of news articles ($r = 0.75$, $p < 0.001$, Fig. 2a) and medical research and patient resources ($r = 0.63$, $p < 0.001$, Fig. 2b).

The search phrase with the highest readability was “rectal cancer covid” at a median grade reading level of 16.5 (IQR 13.75–17), whereas the search phrase with the lowest readability was “breast cancer covid” at a median grade reading level of 12 (IQR 11–15) (Table 1). However, the readability levels among the fourteen most common cancer types were not significantly different amongst cancer subtypes (ANOVA $F = 1.083$, $p = 0.38$).

The median DISCERN scores for the 170 sources assessed was 55 (IQR 51–60) out of 75, which corresponds to “good” quality OHI. The disease site with the highest DISCERN score was “leukemia covid” with a median DISCERN score of 62 (IQR 61–64), whereas the disease site with the lowest DISCERN score was “prostate cancer covid” with a median DISCERN score of 51 (IQR 50–53). Stratified by OHI classification, the collected median DISCERN scores were 54 (IQR 49.5–61.5) for news, 57 (IQR 54–61) for medical sources for research or provider, and 55 (IQR 49–60) for medical sources for patients (Table 1).

With regard to HONcode accreditation, there were 59 HON-accredited websites in English (34.9%), 65 in Spanish (39.9%), and 61 in French (38.6%). DISCERN scores were significantly different when comparing HON-accredited vs. non-HON-accredited sources with median scores of 57 and 54 respectively (*t*-test $p = 0.02$). Reading level was also significantly different when comparing HON-accredited vs. non-HON-accredited sources with median scores of 15 and 13 respectively (*t*-test $p = 0.001$).

Discussion

The elevated risk of poor COVID-19 outcomes faced by cancer patients and the often misunderstood value of physical distancing, mask wearing, and vaccines underscore the importance of access to high-quality OHI, particularly during the COVID-19 pandemic [1]. We found that for 14 cancer types, OHI quality as it pertains to cancer and COVID-19 was on average “good,” although the required grade level readability was well above the AMA recommended sixth-grade level. The recorded data displayed a median readability of 13, requiring at least a high school degree to fully comprehend crucial COVID-19 information, thus potentially exacerbating the disparities in health literacy among cancer patients in America. Furthermore, these findings are likely representative of where most cancer patients are obtaining their information as it pertains to their treatment options amidst the COVID-19 pandemic, given most patients are unlikely to search beyond the first two pages of Google [9].

Unfortunately, little progress has been made towards improving the readability of cancer OHI; studies from the early 2000s showed that readability requirements of cancer OHI at the time were well beyond recommended levels [15]. However, in light of the COVID-19 pandemic and the increased vulnerability of patients with cancer at this time, high-quality and accessible OHI has found new significance [3]. The political and economic valence of the COVID-19 pandemic highlights further the importance of patient education, given the growing proportion of cancer patients accessing OHI [2], and thus a greater need for more widely accessible OHI to educate people with cancer, particularly those receiving treatment during the pandemic. However, the most frequently read COVID-19 and cancer OHI have been found to be significantly above the grade reading level recommended by the AMA, possibly impairing patients’ understanding of treatment options for cancer patients during the coronavirus pandemic. Additionally, the lack of significant difference in readability across cancer types underscores the need to make cancer OHI more readable across the spectrum of disease subtypes.

We further found that higher quality OHI was associated with a greater required reading level; these findings are concerning, as poor-quality OHI may worsen, rather than alleviate, disparities in patient understanding and health literacy. This observed correlation in the discrepancy of accessibility and quality of cancer OHI illuminates the difficulties and limitations for patients with cancer who may experience difficulty comprehending OHI at higher-grade reading levels, and may also influence widespread misunderstanding of treatment options and proper

COVID-19 safety precautions. High required readability was found not just in medical journals or sources intended for physicians, but was also present in OHI designated for patient education and news articles, raising the possibility that high readability of commonly accessed OHI may exacerbate disparities in health literacy specifically for patients with cancer during the pandemic [6].

Though the quality of the recorded sources remains high among OHI for providers and patients based on the DISCERN score, our findings highlight the urgent need for efforts to make OHI more accessible for patients with cancer in line with the American Medical Association’s suggested grade reading level. Efforts devoted to decreasing the required readability of cancer OHI, or providing OHI at the appropriate reading level, may narrow disparities in patient understanding and permit wider audiences of readers to fully comprehend cancer OHI. With the growing importance of OHI in guiding patient treatment decisions and given the impact of the pandemic on cancer care [3], there is great urgency in ensuring that all patients with cancer are accessing and, critically, understanding cancer OHI such that they can make informed treatment decisions.

It is important to note that the observed quality of news sources was lowest in comparison to the other subgroups of sources, as determined by their median DISCERN score, suggesting that news outlets may release less reliable health information as they decrease the grade reading level. This finding has been noted in prior cancer literature as well [16], underscoring the need for more judicious use of news outlets in guiding medical decisions. Educating patients about the quality of various OHI sources—and guiding them towards decisions best suited to their goals and values—is critical.

As patients with cancer refer to the internet for medical information, the utility of OHI is dependent upon appropriate usage by its consumers. eHealth literacy is defined as “the ability to seek, find, and understand health information from electronic sources and apply the knowledge gained to addressing a health problem” [17] and is becoming increasingly crucial in evaluating health literacy and reducing patient costs [6]. The increase in reliance upon OHI for patients with cancer and the accompanying benefits are dependent upon the retention and comprehension of the presented materials by the patient, thus highlighting the importance of eHealth literacy. Therefore, greater efforts should be made to assess and enhance eHealth literacy in conjunction with improving the accessibility of OHI to avoid the spread of misinformation and allow for high-quality OHI to become more prevalent and trusted as a tool to aid informed decision-making.

Our findings are limited by sample size and the subjective nature of validated OHI quality metrics. Additionally, the readability metrics utilized are only capable of approximating patient experience and are subject to variability among

patients with cancer. However, the methodology of the usage of quality and readability metrics was standardized across all 340 sources, resulting in consistent analysis of each source and observable trends of cancer OHI. Furthermore, efforts are needed to make high-quality OHI available at the appropriate reading level for patients with cancer; such efforts are of particular importance during the pandemic and may serve to alleviate disparities in access to healthcare information.

Further investigations are needed to assess the quality and readability of COVID-19 and cancer OHI for other languages given the global implications of the COVID-19 pandemic. Additionally, further studies examining cancer COVID-19 OHI utilizing patient questionnaires could more accurately assess actual patient experiences [2] in terms of their perceived readability and accessibility of these OHI.

Declarations

Conflict of Interest ECD is funded in part through the NIH/NCI Support Grant P30 CA008748 outside the submitted work. All other authors declare no competing interests.

References

1. Yu J, Ouyang W, Chua MLK, Xie C (2020) SARS-CoV-2 transmission in patients with cancer at a tertiary care hospital in Wuhan, China. *JAMA Oncol* 6:1108–1110. <https://doi.org/10.1001/jamaoncol.2020.0980>
2. Dee EC, Muralidhar V, Butler SS, Yu Z, Sha ST, Mahal BA, Nguyen PL, Sanford NN (2020) General and health-related internet use among cancer survivors in the United States: a 2013–2018 cross-sectional analysis. *J Natl Compr Canc Netw* 18:1468–1475. <https://doi.org/10.6004/jnccn.2020.7591>
3. The Lancet Oncology (2021) COVID-19 and cancer: 1 year on. *Lancet Oncol* 22:411. [https://doi.org/10.1016/S1470-2045\(21\)00148-0](https://doi.org/10.1016/S1470-2045(21)00148-0)
4. Fefer M, Lamb CC, Shen AH, Clardy P, Muralidhar V, Devlin PM, Dee EC (2020) Multilingual analysis of the quality and readability of online health information on the adverse effects of breast cancer treatments. *JAMA Surg* 155:781–784. <https://doi.org/10.1001/jamasurg.2020.1668>
5. Perni S, Rooney MK, Horowitz DP, Golden DW, McCall AR, Einstein AJ, Jagsi R (2019) Assessment of use, specificity, and readability of written clinical informed consent forms for patients with cancer undergoing radiotherapy. *JAMA Oncol* 5:e190260. <https://doi.org/10.1001/jamaoncol.2019.0260>
6. Hesse BW, Greenberg AJ, Rutten LJF (2016) The role of Internet resources in clinical oncology: promises and challenges. *Nat Rev Clin Oncol* 13:767–776. <https://doi.org/10.1038/nrclinonc.2016.78>
7. Kripalani S, Weiss BD (2006) Teaching about health literacy and clear communication. *J Gen Intern Med* 21:888–890. <https://doi.org/10.1111/j.1525-1497.2006.00543.x>
8. Schwartz K, Northrup J, Israel N, Crowell K, Lauder N, Neale AV (2003) Use of on-line evidence-based resources at the point of care. *Fam Med* 35:251–256
9. Morahan-Martin JM (2004) How internet users find, evaluate, and use online health information: a cross-cultural review. *Cyberpsychol Behav* 7:497–510. <https://doi.org/10.1089/cpb.2004.7.497>
10. Sha ST, Perni S, Muralidhar V, Mahal BA, Sanford NN, Nguyen PL, Dee EC (2020) Trends, quality, and readability of online health resources on proton radiation therapy. *Int J Radiat Oncol Biol Phys* 107:33–38. <https://doi.org/10.1016/j.ijrobp.2019.12.043>
11. Charnock D, Shepperd S, Needham G, Gann R (1999) DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 53:105–111. <https://doi.org/10.1136/jech.53.2.105>
12. Varady NH, Dee EC, Katz JN (2018) International assessment on quality and content of internet information on osteoarthritis. *Osteoarthr Cartil* 26:1017–1026. <https://doi.org/10.1016/j.joca.2018.04.017>
13. Dee EC, Varady NH, Katz JN, Buchmiller TL (2019) Disparity in online health information in pediatric vs. adult surgical conditions. *Pediatr Surg Int* 35:813–821. <https://doi.org/10.1007/s00383-019-04451-y>
14. Dee EC, Varady NH (2020) Radiation oncology online: quality, strategies, and disparities. *J Canc Educ* 35:988–996. <https://doi.org/10.1007/s13187-019-01553-y>
15. Friedman D, Hoffman-Goetz L, Arocha J (2004) Readability of cancer information on the internet. *J Cancer Educ* 19:117–122. https://doi.org/10.1207/s15430154jce1902_13
16. Gantz W, Wang Z (2009) Coverage of cancer in local television news. *J Cancer Educ* 24:65–72. <https://doi.org/10.1080/08858190802664727>
17. Norman CD, Skinner HA (2006) eHealth literacy: essential skills for consumer health in a networked world. *J Med Internet Res* 8:e9. <https://doi.org/10.2196/jmir.8.2.e9>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.