

Original Article

Evaluation of the Quality and Readability of Internet-Based Health Information Related to COVID-19

Zeynep Gungormus, PhD

Professor, Department of Public Health Nursing, Faculty of Health Sciences, Gaziantep Islamic Science and Technology University, Gaziantep, Turkey

Metin Gungormus, PhD

Professor, Department of Surgery, Faculty of Dentistry, Gaziantep University, Gaziantep, Turkey

Correspondence: Zeynep Gungormus, PhD Professor, Department of Public Health Nursing, Faculty of Health Sciences, Gaziantep Islamic Science and Technology University, Gaziantep, Turkey
E-mail: gungormusz@yahoo.co.uk

Abstract

Objective: The aim of this study was to evaluate the quality and readability of web-based health information on COVID-19.

Material and Methods: An internet search through Google search engine was done with eight search terms related to COVID-19. The first 100 websites were included in the final analysis. The quality of websites was assessed using the DISCERN Instrument, the Journal of the American Medical Association (JAMA) benchmarks, the Ensuring Quality Information for Patients (EQIP) Scale and the Information Quality Tool (IQT). The readability of websites was assessed using the Flesch Reading Ease (FRES) Formula, the Flesch-Kinkaid Reading Grade Level (FKRGL), the Gunning Frequency of Gobbledygook (FOG), the SMOG Index (SMOG), the Coleman-Liau Index (CLI) and the Automated Readability Index (ARI).

Results: The mean DISCERN, JAMA benchmarks, EQIP, and IQT scores were 31.71, 0.59, 42.71 and 5.52, respectively and these findings showed that the websites screened were low quality. The mean scores for FRES, FKRGL, FOG, SMOG, CLI and ARI scores were 52.26, 10.49, 12.78, 9.45, 13.40 and 11.93, respectively. These readability tools showed that the websites screened were difficult or very difficult to read.

Conclusions: As a result, in this study, the quality and readability of websites concerning COVID-19 were evaluated by various validation tools, and it was determined that online health information on COVID-19 were of suboptimal.

Keywords: Quality, readability, web-based health information, COVID-19, Coronavirus disease.

Introduction

Coronavirus disease 2019 (COVID-19) is a new infectious disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (Yilmaz et al., 2021). COVID-19 was first detected in Wuhan, China in late 2019, and the virus quickly spread across the world (Miller et al., 2020). On March 11, 2020, the World Health Organization (WHO) declared COVID-19 outbreak to be a global pandemic (WHO Director-General's opening remarks at

the media briefing on COVID-19, 2020). As of December 21, 2020, there have been over 75 million cases worldwide and more than 1700000 people have died from the virus (WHO Weekly operational update on COVID-19, 2020). But coronavirus cases are still soaring and spreading around the world, and every day, we hear new numbers of cases and deaths. This kind of global pandemic elicits various reactions such as stress, anxiety, fear, and panic from people (Dubey et al., 2020).

The high rate of transmission, the rapid spread, dangerous complications, uncertainties about the disease, the rapid increase in the number of cases, and lack of specific treatments cause fear and panic among people and they have raised the demand for information about the disease. The most common used source of information for the public to follow and stay updated on health information is internet. However, not much is known about the quality and readability of COVID-19 related online health information, and it is also suggested that insufficient information and/or misinform has the potential to influence decisions made by individuals, and they could contribute to unnecessary public panic and result in undesirable responses. For this reason, it is important that the information on the internet is reliable, accurate, high quality, readable and understandable to prevent the panic and the undesirable responses associated with the disease. The aim of this study was to evaluate the quality and readability of online health information related to COVID-19.

Materials and Methods

Selection of Websites: Google search engine was used to identify websites. Eight search term “COVID-19 causes,” “COVID-19 complications”, “COVID-19 treatment”, “COVID-19 vaccine”, “COVID-19 symptoms”, “COVID-19 transmission”, “COVID-19 outbreak”, “COVID-19 prevention” were used. Websites containing irrelevant content, duplicate websites, websites requiring a payment to view the content, scientific articles were excluded, and the first 100 websites meeting inclusion and exclusion criteria were included.

Quality Assessment: The quality of websites was assessed using the DISCERN instrument, the Journal of the American Medical Association (JAMA) benchmarks, the Ensuring Quality Information for Patients (EQIP) Scale and the Information Quality Tool (IQT).

DISCERN Instrument: The DISCERN instrument is a reliable and valid tool for assessing the quality of written health information. The DISCERN consists of 16

questions, and these questions are categorized into three sections. The first section assesses “reliability” of the publication; the second section evaluates the “quality” of information about treatment choices; the third section evaluates “overall quality” of the publication (Charnock et al., 1999).

The total DISCERN score is ranges between 15 and 75, and low scores indicate poor quality, high scores good quality. The total quality each websites was classified as high (≥ 65 points), moderate (33–64 points), or low (16–32 points).

Ensuring Quality Information for Patients (EQIP): EQIP is a 20-item tool and used to assess the reliability, validity and utility of written health information. The total EQIP score ranges from 0% to 100% and low scores indicate poor quality, and high scores indicate good quality (Moult et al., 2004).

Journal of American Medical Association (JAMA) Benchmarks: The JAMA benchmarks are used to evaluate the quality of healthcare websites, and consist of four quality measures: “Authorship”, “Attribution”, “Disclosure”, “Currency”. Each item requires a yes (1 point) or no (0 point) answer. The total JAMA score ranges between 0 and 4 (Silberg et al., 1997).

Information Quality Tool (IQT): IQT is a 21-item tool and used to evaluate the quality of information on the Internet. This scale includes items relation to “authorship,” “sponsorship,” “currency,” “accuracy,” “confidentiality” and “navigability”. Each item requires a yes (1 point) or no (0 point) answer. Total IQT score is varied 0 to 21 and low scores indicate poor quality, high scores good quality (Ademiluyi et al., 2003; Irwin et al., 2011).

Readability Assessment: The readability levels of websites were assessed using the Flesch Reading Ease Formula (FRES), the Flesch-Kinkaid Reading Grade Level (FKRGL), the Gunning Frequency of Gobbledygook (FOG), the SMOG Index (SMOG), the Coleman-Liau Index (CLI) and the Automated Readability Index (ARI) (Kher et al., 2017; Jayaretne et al., 2014).

The Readability scores were calculated automatically with an online Readability

Calculator (<https://www.webpagefx.com>). FRES indicates the readability of the texts, and the other readability tools are related to the educational level of the individual and estimate the years of education the reader requires to understand the text.

FRES score is categorized as very difficult (scores 0-29), difficult (30-49), fairly difficult (50-59), standard (60-69), fairly easy (70-79), easy (80-89) and very easy (90-100).

FKRGL scale is categorized as easy (≤ 6 th-grade level) or difficult (≥ 10 th-grade level) to read. The ideal FOG index score is 7 or 8, with a score above 12 accepted as very difficult for most people.

Results

The results related to quality of online health information on COVID-19 are presented Table 1. The mean DISCERN score was 31.710 (5.340) (range, 21-44), and the DISCERN instrument showed that 43 % of the websites had moderate quality and 57 % showed low quality.

The mean JAMA was 0.590 (0.87) (range, 0-4), and 61% of the websites did not fit any of the JAMA benchmark criteria, 24 % achieved only 1 criterion, 11 % achieved 2 criteria, 3 % achieved 3 criteria, and 1 % achieved 4 criteria. The mean IQT was 5.520 (3.398) (range, 2-17) and the mean EQIP was 42.710 (9.445) (range, 25-67).

The mean readability values for each readability scales are presented in Table 2. The mean readability grade score of the websites was significantly higher than the recommend sixth-grade reading level. The mean FRES score was 52.265 (SD, 10.852) (range, 0.9-74), which is classified as “fairly difficult” to read. Only three websites (3%) had a readability score of above 70 (Fairly easy). Readability level of 18% of web sites was “standard”, and that of the other web sites was “fairly difficult” or “difficult”.

The FKRGL ranged from 1.9 to 22.2, with an average of 10.497 (2.198), which is considered “difficult to read.” Only 3 article (3%) had an acceptable FKRGL score, and readability level of the other web sites was “difficult”.

The mean FOG was 12.787 (2.225) (range, 7.7-24.5), the mean SMOG was 9.451 (1.639) (range, 5.2-18), the mean CLI was 13.404 (1.934) (range, 9.6-21.4) and the mean ARI was 11.930 (8.723) (range, 9.6-103), all of which are significantly higher than the recommended sixth-grade level.

Table 1. The results related to quality of online health information on COVID-19.

Instruments	Mean	SD	Minimum	Maximum
DISCERN Instrument				
Reliability	20.12	3.358	13	32
Quality	11.51	3.271	7	22
Overall	3.350	0.880	2	5
Total	31.630	5.340	21	44

EQIP	42.710	9.445	25	67
JAMA				
Authorship	0.210	0.409	0	1
Attribution	0.010	0.100	0	1
Disclosure	0.200	0.402	0	1
Currency	0.170	0.378	0	4
Total	0.590	0.877	0	4
IQT				
Authorship	0.730	1.607	0	6
Sponsorship	0.160	0.368	0	1
Currency	2.490	0.703	2	4
Accuracy	1.590	1.190	0	3
Confidentiality	0.010	0.100	0	1
Navigability	0.540	0.610	0	3
Total	5.520	3.398	2	17

EQIP; Ensuring Quality Information for Patients, JAMA; Journal of American Medical Association benchmarks, IQT; Information Quality Tool

Table 2. The results related to readability of websites.

	Mean	SD	Minimum	Maximum
FRES	52.265	10.852	0.900	74
FKRGL	10.497	2.198	1.900	22.200
FOG	12.787	2.225	7.700	24.500
SMOG Index	9.451	1.639	5.200	18.00
CLI	13.404	1.934	9.600	21.400
ARI	11.930	8.723	9.600	103.00

FRES; Flesch Reading Ease Formula, FKRGL; Flesch-Kinkaid Reading Grade Level, FOG; Gunning Frequency of Gobbledygook, CLI; Coleman-Liau Index, ARI; Automated Readability Index

Discussion

The COVID-19 outbreak is a major global pandemic that requires adequate and effective communication. The best way to do so is to create resources that are written at the recommended reading level, and are understandable, and are high-quality, reliable, accurate (Kruse et al., 2020). The internet has become one of the most popular resources for finding health information for people, especially during the COVID-19 outbreak. However, it is known that readability and quality of the websites related to the health information are relatively poor. On the other hand, it has been suggest that misinformation is major concern during this pandemic for people, and nowadays, any false and widely circulated information can trigger mass panic, hysteria and adversely impact individual health and wellness (Jayasinghe et al., 2020). The best way to prevent this is to be well informed about the COVID-19 virus, the disease it causes, how it spreads and its causes, complications, treatment, and symptoms.

This study was conducted to evaluate the quality and readability of web-based information related to COVID-19, and it was determined that the quality and readability of web-based health information on COVID-19 were to be of suboptimal quality. Kruse et al performed a study on web-based patient education materials on COVID-19 from US academic medical centers in June 2020, and they reported that the readability, content, and quality of web-based patient education materials addressing COVID-19 was lacking (Kruse et al., 2020).

However, in our study, the mean DISCERN score was 31.71, indicating poor quality of information. The DISCERN instrument showed that 43 % of the websites had moderate quality and 57 % showed low quality. The mean JAMA was 0.59 and 61% of the websites did not fit any of the JAMA benchmark criteria, and only 1 % achieved 4

criteria. In addition, the average IQT was 5.52 and the average EQIP was 42.7. These scores show that the websites related to COVID-19 are poor-quality and the need to begin review of the sites and replacement. These results were consistent with the results of other studies. In a study by Fan et al on March 2020, authors determined that the mean total score for DISCERN, JAMA and EQIP were 38.00, 2.69 and 17.78, respectively, and most websites were inadequate (Fan et al., 2020). In a study by Cuan-Baltazar et al using the search terms "Coronavirus" and "Wuhan" in February 2020, they reported that 39.1% of the websites did not have any of the categories required by the JAMA, and only 10.0% of the websites had the four quality criteria required by JAMA. In addition, they determined that 70.0% of the websites had a low DISCERN score and none of them had a high score (Cuan-Baltazar et al., 2020).

In a study by Jayasinghe et al in the first week of May 2020, they shown the quality, readability, usability, and reliability of the information on COVID-19 on majority of websites providing health information to the general public were to be of substandard quality. In their study, authors determined that the median DISCERN score was 49.5 (fair quality), the median FRES was 54.1 (fairly difficult to read), and only 3.6 % of websites had a readability score of above 70 which is the recommended standard (Jayasinghe et al., 2020).

The National Institutes of Health and the US Department of Health and Human Services recommend that Internet-based patient education materials (IPEMs) be written at or below the seven-grade level (Patel et al., 2020). In the present study, the mean readability grade score of all websites was significantly higher than the recommend sixth-grade reading level. The mean FRES score was 52.26, which is classified as "fairly difficult" to read. Only three websites (3%) had a

readability score of above 70. Eighteen of 100 web sites (18%) had readability indices that were categorized as “standard” to read, and the other web sites had readability indices that were categorized as “fairly difficult” to read or difficult. The FKRGL ranged from 1.9 to 22.2, with an average of 10.49, which is considered “difficult to read.” Only 3 article (3%) had an acceptable FKRGL score, and the other web sites had readability indices that were categorized as “difficult” to read. In addition, the mean FOG was 12.78, the mean SMOG was 9.45, the mean CLI was 13.40 and the mean ARI was 11.93, all of which are significantly higher than the recommended sixth-grade level. Our findings are consistent with those of previous studies.

Szmuda et al determined that the mean FRES, FKRGL, FOG, SMOG Index and CLI scores were 44.14, 12.04, 14.27, 10.71 and 12.69, respectively, and online educational articles on COVID-19 provided information too difficult to read for the general population, and they noted the readability of articles regarding COVID-19 need to improve (Szmuda et al., 2020). Similarly, Worrall et al evaluated the readability of online information relating to COVID-19 in Ireland, the United Kingdom, Canada and the United States, and they determined that the majority of web pages relating to COVID-19 were not at a universal reading level in these four major English-speaking regions (Worrall et al., 2020).

Conclusion: As a result, in this study, the quality and readability of websites concerning COVID-19 were evaluated by various validation tools, and it was determined that online health information on COVID-19 were of suboptimal. When unreliable or false information is spread, the grave danger or risk is that information which is based on truth ends up having only marginal impact. Therefore, the websites related to COVID-19 should be constantly audited, and updated according to the developments.

References

- Ademiluyi G, Rees CE, Sheard CE. (2003). Evaluating the reliability and validity of three tools to assess the quality of health information on the Internet. *Patient Educ Couns*; 50(2):151-5. [https://doi.org/10.1016/S0738-3991\(02\)00124-6](https://doi.org/10.1016/S0738-3991(02)00124-6)
- 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(2):105-11. <http://dx.doi.org/10.1136/jech.53.2.105>
- Cuan-Baltazar JY, Muñoz-Perez MJ, Robledo-Vega C, Pérez-Zepeda MF, Soto-Vega E. (2020). Misinformation of COVID-19 on the Internet: Infodemiology Study. *JMIR Public Health Surveill*; 6(2):e18444. <https://preprints.jmir.org/preprint/18444>
- Dubey S, Biswas P, Ghosh R, Chatterjee S, Dubey MJ, Chatterjee S, Lahiri D, Lavie CJ. (2020). Psychosocial impact of COVID-19. *Diabetes Metab Syndr*; 14(5):779-788. <https://doi.org/10.1016/j.dsx.2020.05.035>
- Fan KS, Ghani SA, Machairas N, Lenti L, Fan KH, Richardson D, Scott A, Raptis DA. (2020). COVID-19 prevention and treatment information on the internet: a systematic analysis and quality assessment. *BMJ Open*; 10(9):e040487. <http://dx.doi.org/10.1136/bmjopen-2020-040487>
- Jayarathne YS, Anderson NK, Zwahlen RA. (2014). Readability of websites containing information on dental implants. *Clin Oral Implants Res*; 25(12):1319-24. <https://doi.org/10.1111/clr.12285>
- Jayasinghe R, Ranasinghe S, Jayarajah U, Seneviratne S. (2020). Quality of online information for the general public on COVID-19. *Patient Educ Couns*; 103(12):2594-7. <https://doi.org/10.1016/j.pec.2020.08.001>
- Irwin JY, Thyvalikakath T, Spallek H, Wali T, Kerr AR, Schleyer T. (2011). English and Spanish oral cancer information on the internet: a pilot surface quality and content evaluation of oral cancer web sites. *J Public Health Dent*; 71(2):106-16. <https://doi.org/10.1111/j.1752-7325.2010.00207.x>
- Kher A, Johnson S, Griffith R. (2017). Readability Assessment of Online Patient Education Material on Congestive Heart Failure. *Adv Prev Med*; 2017:9780317. <https://doi.org/10.1155/2017/9780317>
- Kruse J, Toledo P, Belton TB, Testani EJ, Evans CT, Grobman WA, Miller ES, Lange EMS. (2020). Readability, content, and quality of COVID-19 patient education materials from academic medical centers in the United States.

- Am J Infect Control; 28:S0196-6553(20)31031-2. <https://doi.org/10.1016/j.ajic.2020.11.023>
- Miller HV, Ripepi M, Ernstes AM, Peguero AA. (2020). Immigration Policy and Justice in the Era of COVID-19. *Am J Crim Justice*; 11:1-17. <https://doi.org/10.1007/s12103-020-09544-2>
- Moult B, Franck LS, Brady H. (2004). Ensuring quality information for patients: development and preliminary validation of a new instrument to improve the quality of written health care information. *Health Expect*; 7:165-75. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1369-7625.2004.00273.x>
- Patel CR, Sanghvi S, Cherla DV, Baredes S, Eloy JA. (2015). Readability Assessment of Internet-Based Patient Education Materials Related to Parathyroid Surgery. *Ann Otol Rhinol Laryngol*; 124(7):523-7. <https://doi.org/10.1177/0003489414567938>
- Silberg WM, Lundberg GD, Musacchio RA. (1997). Assessing, controlling, and assuring the quality of medical information on the Internet. *JAMA*; 277:1244-5. doi:10.1001/jama.1997.03540390074039
- Szmuda T, Ozdemir C, Ali S, Singh A, Syed MT, Słoniewski P. Readability of online patient education material for the novel coronavirus disease (COVID-19): a cross-sectional health literacy study. *Public Health*. 2020;185:21-25. <https://doi.org/10.1016/j.puhe.2020.05.041>
- Weekly operational update on COVID-19 - 21 December 2020. <https://www.who.int/publications/m/item/weekly-operational-update-on-COVID-19---21-december-2020>.
- WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-COVID-19---11-march-2020>
- Worrall AP, Connolly MJ, O'Neill A, O'Doherty M, Thornton KP, McNally C, McConkey SJ, de Barra E. (2020). Readability of online COVID-19 health information: a comparison between four English speaking countries. *BMC Public Health*; 20(1):1635. <https://doi.org/10.1186/s12889-020-09710-5>
- Yilmaz A, Marklund E, Andersson M, Nilsson S, Andersson LM, Lindh M, Gisslén M. (2021). Upper Respiratory Tract Levels of Severe Acute Respiratory Syndrome Coronavirus 2 RNA and Duration of Viral RNA Shedding Do Not Differ Between Patients With Mild and Severe/Critical Coronavirus Disease 2019. *J Infect Dis*; 223(1):15-18. <https://doi.org/10.1093/infdis/jiaa632>