ORIGINAL ARTICLE

Quality and readability of online information on ankylosing spondylitis

Burhan Fatih Kocyigit[1] · Tuba Tulay Koca[1] · Mazlum Serdar Akaltun[2]

Abstract

Introduction/objective Obtaining online health-related information is becoming increasingly popular among patients. The attainment of information through websites is easy and practical, but there is no mechanism to check the accuracy and quality of this information. This leads to concerns about information from websites. Therefore, we aimed to evaluate the quality and readability of ankylosing spondylitis–related websites in this study.

Methods This is a descriptive study. Websites were searched on a popular search engine with the search term *ankylosing spondylitis* on March 2, 2019. We recorded the URLs of the first 200 websites listed in the query results. Typologies, quality, and readability were evaluated on these websites. Websites were divided into eight categories (commercial, government, health portal, news, non-profit, professional, scientific journal, and others) according to typology. The JAMA scoring system and the presence of HONcode certification were used to assess the quality. The Flesch-Kincaid grade and the Simple Measure of Gobbledygook were used to evaluate the readability.

Results Of the websites analyzed, 46% were in the high-quality group. We found that scientific journals and news were of higher quality, and commercial and other websites were of poorer quality. The average readability grades of the websites were 8.59 ± 2.42 and 7.33 ± 1.54, which were slightly worse than the recommended value. Additionally, the readability grades were significantly higher on high-quality websites ($p < 0.05$).

Conclusion The quality of information on websites is variable. High-quality information about ankylosing spondylitis is available online, particularly from scientific journals and news. The poor readability of websites that provide high-quality information is a problem for patients with low health literacy. Editors should take into account readability while aiming to present high-quality information on websites.

Key Points
• *Websites have become an important source of health-related information in parallel with the increase in internet use.*
• *Less than half of the ankylosing spondylitis–related websites (46%) were of high quality according to JAMA scores.*
• *The average readability grades of the ankylosing spondylitis–related websites were slightly worse than the recommended values.*
• *High-quality websites had higher readability grades. Therefore, high-quality websites may not be understood by patients with low literacy levels.*
• *No significant difference was found between the websites on the first page (n = 10) and remaining websites (n = 102) in terms of quality and readability.*

✉ Burhan Fatih Kocyigit
bfk2701@hotmail.com

1 Faculty of Medicine, Department of Physical Medicine and Rehabilitation, Kahramanmaraş Sütçü İmam University, Kahramanmaraş, Turkey

2 Department of Physical Medicine and Rehabilitation, Necip Fazıl State Hospital, Kahramanmaraş, Turkey

Introduction

Ankylosing spondylitis (AS) is a systemic inflammatory rheumatic disease whose etiology is complex and not fully understood [1]. Symptoms such as pain, fatigue, sleep disturbance, and mobility limitation have detrimental effects on the quality of life of AS patients [2]. Pharmacologic and non-pharmacologic methods are available to reduce symptoms, slow down the progression of the disease, and treat complications [3].

In parallel with the increase in the use of the Internet in the community, access to information about various diseases, drugs, surgical procedures, and other treatment choices has become easier. It has been shown that about half of the adult American population uses the Internet to obtain health-related information [4]. The main reasons for patients referring to online health-related information sources are to verify information from doctors and to find answers to questions in their heads and seek alternative treatments [5]. While the availability of online health information may be useful in increasing patients' participation in treatment, there are concerns about the quality and accuracy of online health-related information [6, 7]. The absence of a mechanism to assess the accuracy and quality of health-related information provided on websites is the main problem. At this point, the importance of health literacy emerges. Patients with low health literacy skills cannot fully assess the accuracy and quality of online health-related information, which may cause the spreading of incorrect or misleading information.

Today's patient profile tends to search for online information after a specific diagnosis. This is even more important in AS, which is more common in young patients who are more likely to use the Internet as a source of health-related information. The lack of complete recovery, a chronic course with exacerbations, and a long treatment duration may direct patients with AS to alternative sources of information. This leads clinicians to encounter patients who obtain information from the Internet and question diagnosis, follow-up, and treatment choices. The Internet is like a pool of accurate and high-quality information as well as incorrect and misleading information. Access to inaccurate or low-quality information from the Internet can potentially lead to treatment delays and failures. Therefore, patients should be referred to websites that provide high-quality and readable information to prevent these results. We aimed to evaluate websites on AS in terms of information of quality and readability. Additionally, we aimed to determine the website typologies that provide high-quality health-related information. Finally, we evaluated the readability of websites that provided high-quality information.

Material and methods

Data collection

This is a descriptive study. We searched websites on Google (https://www.google.com/) using the search term *ankylosing spondylitis* on March 2, 2019. We chose to use a neutral term to obtain a larger website sample. It was shown that Google is the first search engine preference of more than three-quarters of internet users [8]. Therefore, Google was used as a search engine in this study. Due to the potential effects of earlier internet use, cookies and browser history were cleared before

the study. The first 200 URLs of the websites listed in the search result were recorded. The sample size was consistent with similar studies in the literature [8–10]. Websites that required registration or subscription, inaccessible websites, websites in languages other than English, repetitive websites, and websites that contained no information about AS were excluded from the study. After the application of the exclusion criteria, 112 websites remained for analysis.

Analysis of websites

When evaluating websites, a 3-click rule was performed if no criteria appear on the website homepage. This rule suggests that a user should be able to find information on websites with up to three mouse clicks. The 3-click rule is based on the idea that if users cannot get information in three clicks, they will often leave from the website. Therefore, if no specific characteristic was found within three clicks, a website was awarded 0 points for this criterion in the study [11].

Website typology

Websites were divided into eight categories (*commercial*, *government*, *health portal*, *news*, *non-profit*, *professional*, *scientific journal*, and *others*) independently by two researchers (BFK, TTK) according to typology. Typologies were defined as follows: commercial (websites, which sell products or provide services with the aim of profit); government (websites created, administered, or regulated by an official government agency); health portal (websites which provide health-related information on a variety of topics); news (newspaper or magazine websites created to provide news and information); non-profit (charitable/supportive/educational websites that were not established for the purpose of profit-making); professional (websites created by organizations or individuals with professional medical qualifications); scientific journal (online scientific journals or academic publishing). Websites that could not be classified in any type of typology were evaluated as "others." If there was an inconsistency in the evaluation, a third independent researcher (MSA) evaluated the website and the final decision was made.

The evaluation of the quality of health information was performed using the Journal of American Medline Association (JAMA) score and Health on the Net Foundation certification (HONcode).

JAMA score

The JAMA scoring system includes four parameters: authorship, specifying of date, appropriate references, and ownership [12]. The existence of each parameter is evaluated as 1 point and the absence as 0 points. Thus, the maximum score is 4 and the minimum score is 0. When a website received 3 or 4

points, it was evaluated as high quality [13]. JAMA scores were determined by two independent researchers (BFK, TTK). If there was a discrepancy between the two researchers' JAMA scores, the website was evaluated and finalized by a third independent researcher (MSA). The kappa score of the study was 0.74.

HONcode certification

Websites were examined for the presence of HONcode certification (ref). HON approval stamps were searched for on the websites. HONcode certificates are provided by an independent organization that considers transparency, JAMA scoring, and ethical principles. HONcode certification is a costly process that requires webmasters to apply voluntarily. This feature may limit its use [14, 15].

Readability

The Flesch-Kincaid (FK) grade and the Simple Measure of Gobbledygook (SMOG) were used to evaluate the readability of websites. An online tool (https://www.webfx.com/tools/read-able/) was used to obtain the scores of the two different formulas. The FK evaluates the average sentence length and the average number of syllables per word in the calculation of readability grade. In addition to the parameters in the FK system, the SMOG formula evaluates the number of polysyllabic words in 30 sentences. Higher scores indicate lower readability levels of the websites [9].

Ethics statement

In this study, the researchers did not assess any human participants or animals. Websites that anyone can access were assessed. Therefore, there was no need for the approval of the ethics committee for the study.

Statistical analysis

The statistical analysis of the data was performed using the Statistical Package for the Social Sciences version 20.0 package program (SPSS Inc., Chicago, IL, USA). Number, percentage, mean ± standard deviation, and median were used for the expression of data. Normality of distribution was evaluated using the Shapiro-Wilk test. For categorical variables, comparisons of groups were performed using the chi-square or Fisher's exact tests. Comparisons of two groups were executed using the Mann-Whitney test for continuous variables. Spearmen's rho test was used for the correlation analysis. The significance level was considered as 0.05.

Results

Of the 112 websites, the most frequent typologies were health portal (23.2%), commercial (19.6%), professional (18.8%), and non-profit (16.1%). The distribution of website typologies is presented in Table 1. Internet users typically review websites on the first page of query results [16]. The first page includes 10 websites. The website typologies listed on the first page are presented in Table 1. When we compared the typologies between the first 10 websites and the remaining websites ($n = 102$), no significant difference was found between the groups ($x^2 = 7.812$; $p > 0.05$). This result demonstrates that the distribution of typology did not change between the first 10 websites and the remaining websites ($n = 102$).

Of the 112 websites, 51 (46%) were classified in the high-quality group (JAMA score ≥ 3) and 20 (17.9%) presented HONcode stamps. When the typologies were evaluated in terms of quality, 100% of the scientific journals, 75% of the news sites, 57.1% of the government sites, 52.4% of the professional sites, 50% of the health portals, 27.8% of the non-profit sites, 18.2% of the commercial sites, and 0% of the "other" websites were in the high-quality group. The JAMA scores of the website typologies are demonstrated in Table 2. Additionally, no significant

Table 1 Distribution of the website typologies

Typology	n (%)	Distribution of typologies on the first page, n (%)
Commercial	22 (19.6)	0 (0)
Government	7 (6.3)	1 (10)
Health portal	26 (23.2)	4 (40)
News	8 (7.1)	2 (20)
Non-profit	18 (16.1)	2 (20)
Professional	21 (18.8)	1 (10)
Scientific journal	8 (7.1)	0 (0)
Others	2 (1.8)	0 (0)

n number, % percentage

Table 2 JAMA scores of the website typologies

Typology	JAMA score*
Commercial	1 (1–3)
Government	3 (1–4)
Health portal	2.5 (1–4)
News	3.5 (2–4)
Non-profit	1 (1–4)
Professional	3 (1–4)
Scientific journal	4 (4–4)
Others	1.5 (1–2)

*Data are expressed as median (minimum–maximum)

difference was detected between the first 10 websites (median = 2.5, min = 1, max = 4; IQR = 2.25) and the remaining websites (median = 2, min = 1, max = 4; IQR = 2) in terms of JAMA scores ($p = 0.753$). The main typologies of high-quality websites were scientific journals, news sites, and government websites. Typologies with low quality were "others," commercial, and health portals. In addition, there was no significant difference in terms of quality between the first 10 websites and the remaining websites ($n = 102$).

The mean FK and SMOG grades of the websites were 8.59 ± 2.42 (95% CI 8.13–9.05) and 7.33 ± 1.54 (95% CI 7.04–7.62). No significant difference was found between the websites on the first page ($n = 10$) and the remaining websites ($n = 102$) in terms of FK and SMOG grades ($p = 0.384$ and $p = 0.885$) (Table 3). On the other hand, FK and SMOG grades were significantly higher in high-quality websites ($p = 0.005$ and $p = 0.003$) (Table 4). Additionally, significant and positive correlations were detected between JAMA score and FK and SMOG grades ($r = 0.261$, $p = 0.006$, and $r = 0.220$, $p = 0.02$). These results reveal that there was no significant difference between the top 10 websites and the remaining websites ($n = 102$) in terms of readability. Additionally, high-quality websites had poor readability grades.

Discussion

The Internet is commonly used by patients as a source of health-related information [17]. Free, rapid, and easy access to information is the main reasons for consumers to prefer the Internet. Easy and quick access to health-related online

information has affected patients' decision-making process and has changed the nature of patient-clinician communication/interaction. The passive and physician-dependent structures of patients have changed in obtaining information about health [18]. However, various questions have emerged about obtaining online health-related information:

- Do online information sources provide high-quality information?
- Which website typologies provide health-related information more frequently?
- Which website typologies provide high-quality information?
- Are websites that provide health-related information readable?

In this study, we aimed to find answers to the above-mentioned questions on ankylosing spondylitis–related websites.

The most frequent typologies in this study were health portal, commercial, professional, and non-profit. Higher rates of commercial websites in query results may be considered as a problem. There are concerns that websites with financial purposes may provide biased, incomplete, or incorrect information. Another concern is that search engines can provide high visibility to commercial websites, directing Internet consumers to low-quality information. In contrast to all these concerns, no commercial website was detected on the first page of the query results in this study. Consistent with our results, Chumber et al. [11] and Yaqub et al. [19] detected a high rate of commercial websites in their study. Furthermore, it was reported that there was no commercial website on the first pages of their results. A significant proportion of internet consumers do not review beyond the first page of the query results. Thus, this ranking of Google can be accepted as beneficial when considering possible incorrect or inaccurate information from websites that prioritize commercial gains.

Of the 112 websites, 51 (46%) were classified in the high-quality group and 20 (17.9%) presented HONcode stamps. Arif et al. [9] reported the high-quality website rate as 37%, but the HONcode certificated website rate was found as 7%. Maki et al. [15] detected lower rates of websites with HONcode stamps. On the other hand, Basavakumar et al. [8] reported the high-quality website rate as 43% and websites with HONcode stamps as 53%. Several reasons may lead to

Table 3 Comparison of readability grades between websites on the first page and remaining websites

	Websites on the first page ($n = 10$)	Remaining websites ($n = 102$)	p
FK grade*	7.9 (6.6–10.3; 1.8)	8.5 (2–25.2; 2.25)	0.384
SMOG grade*	7.3 (4.9–8.7; 1.3)	7.2 (1.8–16.3; 1.6)	0.885

n number, *FK* Flesch-Kincaid, *SMOG* Simple Measure of Gobbledygook

*Data are expressed as median (minimum–maximum; interquartile range)

Table 4 Comparison of readability grades between high-quality websites and the remaining websites

	High-quality websites ($n = 51$)	Remaining websites ($n = 61$)	p
FK grade*	8.75 (5.9–25.2; 3.13)	8.2 (2–11.2; 1.65)	0.005
SMOG grade*	7.65 (4.9–16.3; 1.63)	7 (1.8–9.4; 1.45)	0.003

n number, *FK* Flesch-Kincaid, *SMOG* Simple Measure of Gobbledygook

*Data are expressed as median (minimum–maximum; interquartile range)

the relatively low rate of websites with HONcode stamps. HONcode stamps are provided in a process that requires a webmaster to apply voluntarily. The application and evaluation process brings a cost to websites. Therefore, it is not appropriate to consider websites without HONcode stamps as low quality.

The main typologies of high-quality websites were scientific journals, news sites, and government websites. On the other hand, typologies with lower JAMA scores were "others," commercial, and health portals. Different results were reported about website typologies that provide high-quality information. Chumber et al. [11] reported that the high-quality website typologies were health portals and professional sites. Additionally, commercial websites had the lowest JAMA scores. Arif et al. [9] found that scientific journals, news sites, and health portal typologies had higher JAMA scores, and commercial and professional website typologies had lower JAMA scores. The most complete typology was detected as non-profit organizations in another study [8]. We consider that the main reason for the difference between the above-mentioned studies is the evaluation of websites on different diseases such as breast cancer, diabetic neuropathy, and fibromyalgia. Website typologies should be taken into account in order to reach high-quality and accurate health-related information.

No significant difference was found between the websites on the first page ($n = 10$) and the remaining websites ($n = 102$) in terms of JAMA scores and readability. Basavakumar et al. [8] and Arif et al. [9] reported no significant difference between the websites on the first page and the remaining sites in JAMA scores. Readability parameters were found to be slightly higher for websites on the first page [8]. Although Arif et al. [9] found no significant difference in SMOG grades, they found a lower FK grade for websites on the first page. Our results demonstrated that websites had a general trend in terms of JAMA scores and readability parameters, and this did not differ between first-page websites and subsequent pages on websites.

In our study, AS websites had relatively better readability but were above the sixth grade level, which is the recommended level of readability [20]. Additionally, we found poor readability grades on high-quality websites. The average FK grade was found as 12 in studies conducted on Parkinson's disease and lateral epicondylitis [21, 22]. These results are particularly important because patients with less education and lower

health literacy levels have a certain disadvantage. Websites with high readability grades may not be understood by patients or may cause the spreading of misleading/incorrect information, even if they provide accurate information. Website editors should take into account the quality of information, as well as readability to ensure the accessibility and ease of understanding of online consumer–oriented health services.

This study has some limitations. We used only one search term and one search engine. We only evaluated websites that were in English. Websites were reviewed with a single snapshot. The Internet has a dynamic structure and its content is constantly changing. Additionally, the geographic location where website searches are performed may affect results. Although we know that internet consumers are more likely to evaluate websites on the first page, we do not know which of these websites are being reviewed more. In future studies, increasing the number of search terms and engines may be beneficial. Additionally, informative images and relevant videos to support written text should be evaluated. Studies evaluating websites in languages other than English will contribute to the literature. We recommend the development of new quality assessment tools, including detailed reviews of content completeness in future studies.

More than half of the websites did not provide high-quality information about AS. Additionally, the readability grades of high-quality websites were poor. It is obvious that website founders and editors should seek help in the quality and readability of the information they provide. Patients should question online information and increase their health literacy levels. Patients should consider that information from online sources may not be accurate and should share their thoughts and questions with clinicians before criticizing them. Clinicians should recognize that the communication structure with patients has changed and should answer questions in the mind of patients who have acquired online information. The healthcare system should incorporate online information sources. It should refer patients to the right sources to prevent unnecessary use of health services.

Conclusion

The quality of information related to AS on websites was variable. Scientific journals and news sites were of higher quality, and the quality of non-profit, commercial, and "other"

websites was poor. Our results demonstrated that patients with low health literacy may have difficulty understanding AS-related websites. Website editors should focus on readability as well as information quality. Healthcare providers should take greater responsibility for evaluating websites in terms of scientific quality and readability, and should refer patients to more complete sources of online information. We recommend a protocol evaluating websites that provide health-related information before they are accessible online. It would also be beneficial if a committee periodically reviews health-related websites. Scoring systems that assess the quality and reliability of the information provided on health-related websites would be useful. Scores can be presented on the first page of each website so that patients can understand the level of quality and reliability of the information. Guidelines or recommendations for patients may help them to choose wisely which information or advice they will read and potentially follow.

Compliance with ethical standards

Ethical approval   In this study, the researchers did not assess any human participants or animals. Websites that anyone can access were assessed. Therefore, there was no need for the approval of the ethics committee for the study.

Disclosures   None.

References

1. Smith JA (2015) Update on ankylosing spondylitis: current concepts in pathogenesis. Curr Allergy Asthma Rep 15:489. https://doi.org/10.1007/s11882-014-0489-6
2. Dalyan M, Güner A, Tuncer S, Bilgiç A, Arasil T (1999) Disability in ankylosing spondylitis. Disabil Rehabil 21:74–79
3. Zão A, Cantista P (2017) The role of land and aquatic exercise in ankylosing spondylitis: a systematic review. Rheumatol Int 37:1979–1990. https://doi.org/10.1007/s00296-017-3829-8
4. Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL (2015) Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. J Med Internet Res 17(4):e106. https://doi.org/10.2196/jmir.4126
5. Castleton K, Fong T, Wang-Gillam A, Waqar MA, Jeffe DB, Kehlenbrink L, Gao F, Govindan R (2011) A survey of Internet utilization among patients with cancer. Support Care Cancer 19:1183–1190. https://doi.org/10.1007/s00520-010-0935-5
6. Silberg WM, Lundberg GD, Musacchio RA (1997) Assessing, controlling, and assuringthe quality of medical information on the Internet: caveant lector et view or let the reader and viewer beware. JAMA 277:1244–1255
7. Daraz L, Macdermid JC, Wilkins S, Gibson J, Shaw L (2011) The quality of websites addressing fibromyalgia: an assessment of quality and readability using standardised tools. BMJ Open 1(1):e000152. https://doi.org/10.1136/bmjopen-2011-000152
8. Basavakumar D, Flegg M, Eccles J, Ghezzi P (2019) Accuracy, completeness and accessibility of online information on fibromyalgia. Rheumatol Int 39:735–742. https://doi.org/10.1007/s00296-019-04265-0
9. Arif N, Ghezzi P (2018) Quality of online information on breast cancer treatment options. Breast 37:6–12. https://doi.org/10.1016/j.breast.2017.10.004
10. Manley L, Ghezzi P (2018) The quality of online health information on breast augmentation. J Plast Reconstr Aesthet Surg 71:e62–e63. https://doi.org/10.1016/j.bjps.2018.07.023
11. Chumber S, Huber J, Ghezzi P (2015) A methodology to analyze the quality of health information on the internet: the example of diabetic neuropathy. Diabetes Educ 41:95–105. https://doi.org/10.1177/0145721714560772
12. Silberg WM, Lundberg GD, Musacchio RA (1997) Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor–let the reader and viewer beware. JAMA 277:1244–1245
13. Meric F, Bernstam EV, Mirza NQ, Hunt KK, Ames FC, Ross MI, Kuerer HM, Pollock RE, Musen MA, Singletary SE (2002) Breast cancer on the world wide web: cross sectional survey of quality of information and popularity of websites. British Med J 324:577–581
14. Boyer C, Selby M, Scherrer JR, Appel RD (1998) The health on the net code of conduct for medical and health websites. Comput Biol Med 28:603–610
15. Maki A, Evans R, Ghezzi P (2015) Bad news: analysis of the quality of information on influenza prevention returned by Google in English and Italian. Front Immunol 6:616. https://doi.org/10.3389/fimmu.2015.00616
16. Eysenbach G, Kohler C (2002) How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. British Med J 324:573–577
17. Wald HS, Dube CE, Anthony DC (2007) Untangling the Web—the impact of Internet use on health care and the physician-patient relationship. Patient Educ Couns 68:218–224
18. Gardiner R (2008) The transition from 'informed patient' care to 'patient informed' care. Stud Health Technol Inf 137:241–256
19. Yaqub M, Ghezzi P (2015) Adding dimensions to the analysis of the quality of health information of websites returned by Google: cluster analysis identifies patterns of websites according to their classification and the type of intervention described. Front Public Health 3:204. https://doi.org/10.3389/fpubh.2015.00204
20. Brosnan S, Barron E, Sahm LJ (2012) Health literacy and the clozapine patient. Perspect Public Health 132:39–42
21. Dy CJ, Taylor SA, Patel RM, McCarthy MM, Roberts TR, Daluiski A (2012) Does the quality, accuracy, and readability of information about lateral epicondylitis on the internet vary with the search term used? Hand (N Y) 7:420–425. https://doi.org/10.1007/s11552-012-9443-z
22. Fitzsimmons PR, Michael BD, Hulley JL, Scott GO (2010) A readability assessment of online Parkinson's disease information. J R Coll Physicians Edinb 40:292–296. https://doi.org/10.4997/JRCPE.2010.401