

## Original Articles

# Evaluating the readability, quality and reliability of online patient education materials on chronic low back pain

ERKAN OZDURAN, VOLKAN HANCI, YUKSEL ERKIN

### ABSTRACT

**Background.** There are concerns over the reliability and comprehensibility of health-related information on the internet. We analyzed the readability, reliability and quality of online patient education materials obtained from websites associated with chronic low back pain (cLBP).

**Methods.** On 26 April 2022, the term 'cLBP' was used to perform a search on Google, and 95 eligible websites were identified. The Flesch Reading Ease Score (FRES) and Gunning Fog (GFOG) index were used to evaluate the readability. The *Journal of the American Medical Association* (JAMA) score was used to assess the reliability and the Health on the Net Foundation code of conduct (HONcode) was used to assess quality.

**Results.** The mean (SD) FRES was 55.74 (13.57) (very difficult) and the mean (SD) GFOG was 12.76 (2.8) (very difficult) of the websites reviewed. According to the JAMA scores, 28.4% of the websites had a high reliability rating and 33.7% adhered to the HONcode. Websites of different typologies were found to significantly differ in their reliability and the quality scores ( $p < 0.05$ ).

**Conclusion.** The reading ability required for cLBP-related information on the internet was found to be considerably higher than that recommended by the National Health Institute and had low reliability and poor quality. We believe that online information should have readability appropriate for most readers and must have reliable content that is appropriate to educate the public, particularly for websites that provide patient education material.

Natl Med J India 2024;37:124–30

### INTRODUCTION

Low back pain (LBP) is defined as pain, stiffness or muscle

Dokuz Eylül University, Inciraltı mahallesi, Izmir 35030, Turkey  
ERKAN OZDURAN Department of Physical Medicine and Rehabilitation

VOLKAN HANCI, YUKSEL ERKIN Department of Anesthesiology and Reanimation

Correspondence to ERKAN OZDURAN; [erkanozduran@gmail.com](mailto:erkanozduran@gmail.com)

[To cite: Ozduran E, Hanci V, Erkin Y. Evaluating the readability, quality and reliability of online patient education materials on chronic low back pain. *Natl Med J India* 2024;37:124–30. DOI: 10.25259/NMJL\_327\_2022]

© The National Medical Journal of India 2024

tension localized between the lower gluteal fold and the costal border.<sup>1</sup> LBP has been one of the leading causes of disability for the past 30 years and leads to a major burden on healthcare services and causes productivity losses.<sup>2</sup> More than three-quarters of patients with acute LBP (aLBP) have a good prognosis in terms of disability and pain, and return to work in less than a month.<sup>3,4</sup> Itz *et al.*<sup>5</sup> reported that one-third of individuals with LBP recovered within 3 months but two-thirds reported pain even after 1 year. In addition to social, physiological and genetic causes, high levels of disability and pain, and the presence of comorbid conditions, also play a role in the transformation of aLBP into chronic LBP (cLBP).<sup>6</sup> Superficial heat application, massage, acupuncture, spinal manual therapy, pharmacological treatment and multidisciplinary rehabilitation programmes are some of the treatment modalities that can be used for LBP.<sup>7</sup>

The internet has recently become an important source of health-related information. Patients access medical information through internet-based patient education materials (PEMs), soothe their worries and fears, and save time.<sup>8</sup> A study found that 9 of 10 adult Americans used the internet in 2018 and nearly 75% of them researched health-related issues.<sup>9</sup> According to the National Institutes of Health, the US Department of Health and Human Services, and the American Medical Association, PEMs on the internet should be written at a sixth-grade level.<sup>8,9</sup> If the readability of online material on a website exceeds this threshold, it is likely to be difficult to read and comprehend for the typical reader. As a result, it is critical that health-related material on websites be appropriate for the reader and thoroughly assessed before being used. There is increasing use of health-related information from the internet but it is unclear if the information provided is accurate, reliable, of high quality and readable. For this reason, scientific studies have assessed the quality and readability of information for various diseases on the internet.<sup>9,10</sup> A study on the readability and quality of online information about aLBP reported that the quality of information was low and the content was difficult to read.<sup>11</sup>

Patients who have knowledge about the causes, pathophysiology, treatment and prevention of a disease are better able to participate and comply during the disease prevention or treatment procedures. As with other diseases, the level of knowledge of patients about cLBP is low.<sup>12</sup> It is clear that the transfer of reliable, high quality and readable information on this subject to individuals will play an important role in preventing cLBP, or in easy administration of treatment. We aimed to evaluate websites providing PEMs on cLBP for their readability,

quality and reliability. In addition, we also tried to determine website typologies that provided reliable information about cLBP.

#### METHODS

Our study was conducted with the approval of the Non-Interventional Research Ethics Committee (6488-GOA 2021/20-63). Two independent authors (EO, VH) searched the keyword ‘Chronic low back pain’ on Google ([www.google.com.tr](http://www.google.com.tr)) on 26 April 2022. In case of disagreements among the authors during the evaluation of the websites, the final decision was made by a third independent author (YE). Google search engine was used because based on data from December 2021 it led the sector with a market share of 86.2%.<sup>13</sup>

During the website search, cookies and the computer’s browser history were cleared to ensure that the search results were unaffected by prior access. In addition, the study was conducted by logging out of all Google accounts. Completing the search, the first 200 websites’ uniform resource locators (URLs) were recorded, following the methodology of similar research in the literature.<sup>14,15</sup> The top 10 websites on the first page were ranked as the most viewed websites.<sup>16</sup> The study excluded websites with non-English content, websites without information about cLBP, websites that demand registration or subscription, repetitive websites, websites with video or audio recording content but no written content, and websites with journal articles. In addition, graphics, pictures, videos, tables, figures and list formats in the texts, all punctuation marks, URL websites, author information, addresses and phone numbers, as well as references were not included in the evaluation to avoid erroneous results.<sup>17</sup>

During the website evaluation, if an evaluation criterion could not be identified on the homepage, the three-click rule was used,<sup>18</sup> which states that a website user should be able to find any information in three mouse clicks or less. Although this is not a ‘rule’, it is believed that if information cannot be found in three clicks, the users will be unable to complete their task and will leave the site.

#### *Website typology*

Two independent authors classified websites into six categories based on their typology. Typologies were professional (websites created by organizations or individuals with professional medical qualifications), commercial (websites that sell product for profit), non-profit (non-profit educational/charitable/supporting sites), health portals (websites that provide information about health issues), news (news and information created to provide magazine websites or newspaper), government (websites created, regulated or administered by an official government agency).

#### *Journal of American Medical Association (JAMA) Benchmark Criteria*

The *JAMA* benchmark analyzes online information and resources under 4 criteria: authorship, attribution, disclosure and currency. (*JAMA* score 0–4, Authorship [1 point]: Authors and contributors, their affiliations, and relevant credentials should be provided; Attribution [1 point]: References and sources for all content should be listed; Disclosure [1 point]: Conflicts of interest, funding, sponsorship, advertising, support, and video ownership should be fully disclosed; Currency [1 point]: Dates on which the content was posted and updated should be indicated). The *JAMA* score is used to evaluate the accuracy

and reliability of information. The scorer awards 1 point for each criterion in the text, and the final score ranges from 0 to 4. Four points represent the highest reliability and quality.<sup>19</sup> Websites with a *JAMA* score of 1 are considered as partially sufficient data, 2-partially sufficient data, 3-completely sufficient data. Websites with a *JAMA* score of  $\geq 3$  points were considered highly reliable and those with a *JAMA* score of  $\leq 2$  points were considered to have low reliability.<sup>19</sup>

#### *DISCERN criteria*

The DISCERN criteria, a technique for assessing the quality of websites, consists of 16 questions with scores ranging from 1 to 5.<sup>20</sup> The first 8 questions ask about the website’s basic content, such as ‘are the aims clear?’ and ‘were citations used?’ The last 8 questions test treatment knowledge, such as ‘is it clear that there is more than one treatment option?’. Using the DISCERN criteria, two authors independently examined websites. Averaging the data from the two separate authors yielded the final DISCERN score for each website. The final DISCERN score varies from 16 to 80. A score of 63 to 80 represents ‘excellent’, 51 to 62 represents ‘good’, 39 to 50 represents ‘fair’, 28 to 38 represents ‘poor’ and 16 to 27 represents ‘very poor’.<sup>21</sup>

#### *Health on the Net Foundation code of conduct (HONcode) certification*

The Health on the Net Foundation (HON) was founded to promote the efficient transmission and use of reliable and useful health information via the Internet. HONcode was created by HON to help standardize the accuracy of health-related information on the Internet.<sup>22</sup> To meet the HONcode criteria, the content’s date and source should be disclosed, the authors’ qualifications should be specified, the privacy policy should be explained, the patient–physician relationship should be supported rather than replaced, the website’s financing and advertising policy should be specified, and contact information should be explained.<sup>23</sup> HON grants HONcode certificates to websites as an option. HONcode is an affordable and optional certificate. The HONcode certificate is subject to a price and its use is restricted. We checked if the main page or a connected URL had a HONcode stamp.

#### *Readability*

The following readability formulas were used to assess website readability: Flesch reading ease score (FRES), Flesch–Kincaid grade level (FKGL), Simple Measure of Gobbledygook (SMOG), Gunning Fog index (GFOG), Coleman–Liau score (CL), automated readability index (ARI), and Linsear Write (LW) readability formulas from [www.readabilityscore.com](http://www.readabilityscore.com).<sup>24–28</sup>

All websites’ ranking values were calculated and recorded. Microsoft Office Word 2007 (Microsoft Corporation, Redmond, WA) was used to copy and save the texts. Based on the sixth-grade level specified by the American Medical Association and the National Institutes of Health, the average readability level according to all readability formulas was compared.

#### *Popularity and visibility analysis*

Alexa ([www.alexa.com](http://www.alexa.com)) is a popular traffic engine, and it is frequently used to assess area visibility and popularity.<sup>29</sup> It compares the number of times a website has been visited in the past three months to the number of times other websites have been visited. The higher the score, the more popular the site is because of more clicks.

### Content analysis

Websites were assessed based on their typologies to see if they contained any cLBP-related content (aetiology, diagnosis, symptoms, treatment, surgery, exercise, prevention and risk factors).

### Statistical analysis

Data were uploaded to SPSS Windows 25.0 software (SPSS Inc., Chicago, IL). Continuous values are indicated as mean (SD), while frequency variables are given as number (*n*) and percentage (%). Whether the data with continuous values conformed to the normal distribution pattern was evaluated with the Kolmogorov Smirnov and Shapiro Wilk tests. It was determined that the data with continuous values did not fit the normal distribution pattern. For statistical analysis, the Mann–Whitney U or Kruskal Wallis tests were used to compare groups with continuous values such as readability indices and sixth class level. For comparison of frequency variables, the Pearson Chi-square or Fisher exact tests were used. Pearson correlation test was used in correlation analysis.  $p < 0.05$  was considered statistically significant.

### RESULTS

Our study included 200 websites; 105 were eliminated because they did not match the inclusion criteria, and the remaining 95 were evaluated. Commercial (30.5%) and health portal (17.9%) websites were found to be the most common when compared according to their typologies.

On Google's first page, there are ten search results. There was no statistically significant difference between the first 10 search results and the remaining search results when they were analyzed according to their typologies ( $p = 0.12$ ). There was no statistically significant difference between the readability values of the top 10 websites and the readability values of the remaining websites (FRES, GFOG, GFOG, CL, SMOG;  $p > 0.05$ ). A significant result ( $p = 0.045$ ) was obtained when the Alexa values of the first 10 sites were compared to the Alexa values of the remaining sites. The top 10 websites on the first page were more popular in terms of search, viewing and traffic. There was no significant difference between the presence of JAMA reliability ( $p = 0.06$ ), DISCERN quality ( $p = 0.16$ ), HONcode presence ( $p = 0.12$ ) or typologies ( $p = 0.12$ ) of the top 10 and the remaining 85 websites (Table I).

These 95 websites had a mean (SD) JAMA score of 1.97 (0.99), a DISCERN score of 40.65 (19.94) and an Alexa score of 720406.64 (2029785.31) which suggest low reliability (JAMA score  $< 3$ ) and poor quality (DISCERN score 32–48). About a quarter (28.4%) of the websites were highly reliable with a JAMA score of  $\geq 3$ . The text content of the 95 evaluated websites had a mean (SD) FRES of 55.74 (13.57) (very difficult), and the mean GFOG was 12.76 (2.8) (very difficult). The mean (SD) FKGL and SMOG were 10.12 (2.7) and 9.46 (2.10) years of education, respectively, while the mean (SD) CL index was 10.28 (1.96) years and ARI index was 10.32 (3.0) years of education. The readability indices were compared by site typology, and the results indicated no significant difference ( $p > 0.05$ ). A statistically significant difference was found when the readability index averages of 95 websites were compared to the grade 6 reading level ( $p < 0.003$ ) (Table I).

When the top 10 websites were compared to the other websites using content analysis, there was no statistically

significant difference ( $p > 0.05$ ). Only websites with diagnosis ( $p = 0.023$ ) and symptoms ( $p = 0.007$ ) content had a statistically significant difference in their contents according to typology when all 95 websites were assessed (Table II). Commercial websites frequently mentioned diagnosis and symptoms.

There was a significant difference in JAMA reliability scores ( $p < 0.001$ ), DISCERN quality scores ( $p = 0.002$ ), and HONcode ( $p = 0.008$ ) by site typology. This statistical difference can be explained by higher JAMA reliability scores and DISCERN quality scores in government and professional websites. These scores were found to be lower in commercial content websites. Only 32 (33.7%) of all sites had HONcode. The highest number of HONcode were found on health portals (11; Table III).

The FRES, FKGL, SMOG, GFOG, CL, ARI, LW readability formula averages, JAMA and DISCERN scores, and HONcode entities were analyzed with respect to the site rankings. A weak positive correlation was observed between the JAMA and DISCERN scores ( $r = 0.852$ ,  $p < 0.001$ ) and JAMA and HONcode entities ( $r = 0.351$ ,  $p < 0.001$ ). The readability indices had a correlation, however, there was none between the readability scores and the popularity and visibility analysis index (Alexa; Table IV).

### DISCUSSION

We evaluated whether the internet-based PEM related to cLBP is reliable, high quality and readable. We also tried to determine which types of sites provide highly reliable and readable information. We planned to compare the 10 most visited sites on the first page of a Google search with other sites in terms of quality, reliability and readability. Finally, we evaluated the relationship between the readability of the sites and their quality and reliability. In our study, the readability level of the internet-based PEM on cLBP was found to be well above the 6th grade level recommended by the National Institutes of Health. Website content was found to have low reliability and poor quality. It was found that health-portal-sourced websites provided more reliable and high quality information, whereas commercial websites performed poorly in comparison. By looking at the correlation between JAMA, DISCERN and HONcodes, we found that reliable sites also provided high quality information.

The internet has become a resource not only for patients but also for healthcare providers. More than 70% of adults search for health information online and more than 30% try to diagnose a medical problem in themselves or someone they care for.<sup>10</sup> Approximately 90% of young people and adults in the USA currently use the internet, and internet use is increasing day by day across all age groups.<sup>10</sup>

It is known that online health resources provide patients with information on healthier lifestyle choices and regular health check visits.<sup>30</sup> PEMs strengthen communication between patients and doctors, increase patient awareness, and encourage patient-centered care.<sup>31</sup> By evaluating the readability of online PEMs, previous studies determined that the contents required higher reading ability than that of an average American adult, and it was reported that a more suitable language should be used in PEMs.<sup>32</sup> Texts consisting of long and complex sentences can affect the reader's confidence while the reader tries to gain medical information and cause them to give up on reading the text. According to the US Department of Education, the National Literacy Institute, 32 million American adults are illiterate and 68 million Americans have a reading level below that of fifth

TABLE I. All websites' mean results and statistical comparison of text content to 6th grade reading level

Index	Top 10 (n=10) Mean (SD)	Others (n=85) Mean (SD)	Total (n=95) Mean (SD)	Comparison of the first 10 websites and remaining 85 websites according to parameters (p)	Comparison of the 95 websites' according to 6th grade reading level (p)
<i>Readability indexes</i>					
FRES	56.91 (14.65)	55.61 (13.53)	55.74 (13.57)	0.766	0.003
GFOG	12.59 (3.67)	12.78 (2.7)	12.76 (2.8)	0.956	<0.001
FKGL	9.98 (3.23)	10.14 (2.66)	10.12 (2.7)	0.932	<0.001
CL	10.4 (2.27)	10.27 (1.93)	10.28 (1.96)	0.533	<0.001
The SMOG index	9.18 (2.48)	9.49 (2.07)	9.46 (2.10)	0.976	<0.001
ARI	10.5 (3.59)	10.3 (2.95)	10.32 (3)	0.636	<0.001
LW formula	11.61 (4.48)	11.76 (3.46)	11.75 (3.55)	0.785	<0.001
Grade level	10.3 (2.98)	10.4 (2.5)	10.38 (2.54)	0.845	<0.001
<i>Popularity indexes</i>					
Alexa rank	91290.6 (262448.56)	798075.29 (2138153.38)	720406.64 (2029785.31)	0.045	-
JAMA	1.4 (0.84)	2.04 (0.99)	1.97 (0.99)	0.056	-
DISCERN	31.8 (14.61)	41.69 (20.29)	40.65 (19.94)	0.158	-
JAMA	<i>n (%)</i>	<i>n (%)</i>	<i>n (%)</i>	0.169	-
Insufficient data	6 (60)	27 (31.8)	33 (34.7)	-	-
Partially sufficient data	4 (40)	50 (58.8)	54 (56.8)	-	-
Completely sufficient data	0 (0)	8 (9.4)	8 (8.4)	-	-
DISCERN				0.547	-
Very poor	3 (30)	20 (23.5)	23 (24.2)	-	-
Poor	5 (50)	25 (29.4)	30 (31.6)	-	-
Fair	1 (10)	13 (15.3)	14 (14.7)	-	-
Good	1 (10)	21 (24.7)	22 (23.2)	-	-
Excellent	0 (0)	6 (7.1)	6 (6.3)	-	-
HONcode <i>n (%)</i> +	6 (60)	26 (30.6)	32 (33.7)	0.069	-
HONcode <i>n (%)</i> -	4 (40)	59 (69.4)	63 (66.3)	-	-
Typology				0.116	-
Professional	1 (10)	15 (17.6)	16 (16.8)	-	-
Commercial	3 (30)	26 (89.7)	29 (30.5)	-	-
Non-profit	3 (30)	10 (76.9)	13 (13.7)	-	-
Health portal	0 (0)	17 (20)	17 (17.9)	-	-
News	0 (0)	9 (10.6)	9 (9.5)	-	-
Government	3 (30)	8 (9.4)	11 (11.6)	-	-
FRES Flesch reading ease score    FKGL Flesch-Kincaid grade level    SMOG Simple measure of Gobbledygook    GFOG Gunning Fog score    CL Coleman-Liau score    ARI Automated readability index    LW Linsear Write    JAMA <i>Journal of American Medical Association</i> benchmark criteria    HONcode The Health on the Net Foundation code of conduct					

TABLE II. All websites' content analysis by typology

Content description	Professional	Commercial	Non-profit	Health portal	News	Government	Total	p value	
Aetiology	Present	11 (68.8)	23 (79.3)	12 (92.3)	13 (76.5)	3 (33.3)	7 (63.6)	69 (72.6)	0.054
	Absent	5 (31.3)	6 (20.7)	1 (7.7)	4 (23.5)	6 (66.7)	4 (36.4)	26 (27.4)	
Diagnosis	Present	7 (43.8)	13 (44.8)	11 (84.6)	6 (35.3)	1 (11.1)	5 (45.5)	43 (45.3)	0.023
	Absent	9 (56.3)	16 (55.2)	2 (15.4)	11 (64.7)	8 (88.9)	6 (54.5)	52 (54.7)	
Symptoms	Present	14 (87.5)	21 (72.4)	12 (92.3)	12 (70.6)	5 (55.6)	3 (27.3)	67 (70.5)	0.007
	Absent	2 (12.5)	8 (27.6)	1 (7.7)	5 (29.4)	4 (44.4)	8 (72.7)	28 (29.5)	
Treatment	Present	12 (75)	26 (89.7)	11 (84.6)	12 (70.6)	6 (66.7)	10 (90.9)	77 (81.1)	0.415
	Absent	4 (25)	3 (10.3)	2 (15.4)	5 (29.4)	3 (33.3)	1 (9.1)	18 (18.9)	
Surgery	Present	6 (37.5)	16 (55.2)	9 (69.1)	7 (39.2)	2 (22.2)	4 (36.4)	44 (46.3)	0.293
	Absent	10 (62.5)	13 (44.8)	4 (30.8)	10 (58.8)	7 (77.8)	7 (63.6)	51 (53.7)	
Exercise	Present	7 (43.8)	20 (69)	11 (84.6)	10 (58.8)	5 (55.6)	6 (54.5)	59 (62.1)	0.289
	Absent	9 (56.3)	9 (31)	2 (15.4)	7 (41.2)	4 (44.4)	5 (45.5)	36 (37.9)	
Prevention	Present	4 (25)	15 (51.7)	8 (61.5)	6 (35.5)	8 (88.9)	6 (54.5)	40 (42.1)	0.087
	Absent	12 (75)	14 (48.3)	5 (38.5)	11 (64.7)	1 (11.1)	5 (45.5)	55 (57.9)	
Risk factors	Present	7 (43.8)	14 (48.3)	6 (46.2)	6 (35.3)	1 (11.1)	6 (54.5)	40 (42.1)	0.401
	Absent	9 (56.3)	15 (51.7)	7 (53.8)	11 (64.7)	8 (88.9)	9 (45.5)	55 (57.9)	

Figures in parentheses are percentages

TABLE III. Comparison of JAMA, DISCERN scores, HONcode presences and reading levels according to the typologies of the websites

Item	Professional	Commercial	Non-profit	Health portal	News	Government	p value
<i>n</i> (%)	16 (16.8)	29 (30.5)	13 (13.7)	17 (17.9)	9 (9.5)	11 (11.6)	
Mean (SD) JAMA	2.37 (1.02)	1.27 (0.70)	2.30 (0.85)	2.25 (0.94)	2.11 (0.60)	1.90 (0.99)	<0.001
<i>JAMA benchmark criteria</i>							0.001
Insufficient data ( <i>n</i> =33)	2 (12.5)	19 (65.5)	2 (15.4)	3 (17.6)	1 (11.1)	6 (54.5)	
Partially sufficient data ( <i>n</i> =54)	11 (68.8)	10 (34.5)	10 (76.9)	12 (70.6)	8 (88.9)	3 (27.3)	
Completely sufficient data ( <i>n</i> =8)	3 (18.8)	0 (0)	1 (7.7)	2 (11.8)	0 (0)	2 (18.2)	
Mean (SD) DISCERN	47.5 (22.26)	26.41 (13.51)	50.92 (15.59)	52.7 (19.03)	33.55 (14.34)	43.27 (18.63)	<0.001
<i>DISCERN</i>							0.002
Very poor ( <i>n</i> =23)	2 (12.5)	16 (55.2)	0 (0)	2 (11.8)	2 (22.2)	1 (9.1)	
Poor ( <i>n</i> =30)	6 (37.5)	8 (27.6)	4 (30.8)	2 (11.8)	5 (55.6)	5 (45.5)	
Fair ( <i>n</i> =14)	1 (6.3)	4 (13.8)	3 (23.1)	3 (17.6)	1 (11.1)	2 (18.2)	
Good ( <i>n</i> =22)	4 (25)	1 (3.4)	5 (38.5)	9 (52.9)	1 (11.1)	2 (18.2)	
Excellent ( <i>n</i> =6)	3 (18.8)	0 (0)	1 (7.7)	1 (5.9)	0 (0)	1 (9.1)	
<i>HONcode</i>							0.008
Present ( <i>n</i> =32)	5 (31.3)	5 (17.2)	7 (53.8)	11 (64.7)	1 (11.1)	3 (27.3)	
Absent ( <i>n</i> =63)	11 (68.8)	24 (82.8)	6 (46.2)	6 (35.3)	8 (88.9)	8 (72.7)	
<i>Reading level</i>							0.027
Easy to read	1 (6.3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
Fairly easy to read	2 (12.5)	3 (10.3)	1 (7.7)	6 (35.3)	1 (11.1)	3 (27.3)	
Standard/average <i>n</i> (%)	6 (37.5)	10 (34.5)	3 (23.1)	3 (17.6)	2 (22.6)	0 (0)	
Fairly difficult to read <i>n</i> (%)	6 (37.5)	11 (37.9)	2 (15.4)	2 (11.8)	1 (11.1)	1 (9.1)	
Difficult to read <i>n</i> (%)	0 (0)	4 (13.8)	5 (38.5)	6 (35.3)	5 (55.6)	5 (45.5)	
Very difficult to read <i>n</i> (%)	1 (6.3)	1 (3.4)	2 (15.4)	0 (0)	0 (0)	2 (18.2)	
<i>Readers age (years)</i>							0.437
8–9 (Fourth and fifth graders)	1 (6.3)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	
10–11 (Fifth and sixth graders)	0 (0)	1 (3.4)	1 (7.7)	0 (0)	0 (0)	0 (0)	
11–13 (Sixth and seventh graders)	2 (12.5)	0 (0)	0 (0)	2 (11.8)	1 (11.1)	2 (18.2)	
12–14 (Seventh and eighth graders)	2 (12.5)	5 (17.2)	2 (15.4)	5 (29.4)	1 (11.1)	1 (9.1)	
13–15 (Eighth and ninth graders)	2 (12.5)	5 (17.2)	2 (15.4)	1 (5.9)	0 (0)	1 (9.1)	
14–15 (Ninth to tenth graders)	5 (31.3)	7 (24.1)	1 (7.7)	2 (11.8)	1 (11.1)	0 (0)	
15–17 (Tenth to eleventh graders)	2 (12.5)	4 (13.8)	2 (15.4)	2 (11.8)	1 (11.1)	0 (0)	
17–18 (Twelfth graders)	1 (6.3)	4 (13.8)	2 (15.4)	2 (11.8)	2 (22.2)	1 (9.1)	
18–19 (College level entry)	0 (0)	2 (6.9)	1 (7.7)	3 (17.6)	2 (22.2)	2 (18.2)	
21–22 (College level)	0 (0)	0 (0)	0 (0)	0 (0)	1 (11.1)	2 (18.2)	
College graduate	1 (6.3)	1 (3.4)	2 (15.4)	0 (0)	2 (18.2)	6 (6.3)	

Figures in parentheses are percentages unless specified  
Foundation code of conduct

JAMA *Journal of American Medical Association* benchmark criteria

HONcode The Health on the Net

TABLE IV. Correlation relationships between rank and readability formulas, JAMA, DISCERN scores and HONcode

Rank	Alexa rank		JAMA		DISCERN		HONcode	
	r	p	r	p	r	p	r	p
Mean FRES	-0.152	0.151	0.013	0.897	0.019	0.857	0.273	0.007
Mean GFOG	0.151	0.153	-0.072	0.485	-0.047	0.658	-0.315	0.002
Mean FKGL	0.119	0.263	-0.047	0.652	-0.049	0.638	0.305	0.003
Mean CL	0.137	0.196	-0.067	0.516	-0.070	0.516	-0.150	0.148
Mean SMOG index	0.154	0.144	-0.053	0.607	-0.056	0.591	-0.327	0.001
Mean ARI	0.098	0.356	-0.113	0.275	-0.112	0.279	-0.288	0.005
Mean LW formula	0.103	0.332	-0.088	0.396	-0.088	0.396	-0.344	0.001
Grade level	0.133	0.208	-0.085	0.415	-0.090	0.385	-0.321	0.002
Alexa	-	-	-0.321	0.002	-0.348	0.001	-0.252	0.016
JAMA	-0.321	0.002	-	-	0.852	<0.001	0.351	<0.001
DISCERN	-0.348	0.001	0.852	<0.001	-	-	0.358	<0.001
HONcode	-0.252	0.016	0.351	<0.001	0.358	<0.001	-	-

FRES Flesch reading ease score

FKGL Flesch–Kincaid grade level

SMOG Simple measure of Gobbledygook

GFOG Gunning Fog

CL Coleman–Liau

score ARI automated readability index

LW Linsear Write

HONcode The Health on the Net Foundation code of conduct

JAMA *Journal of American*

Medical Association benchmark criteria

graders.<sup>33,34</sup> Considering the increase in health-related information acquisition from the internet, providing more readable information on websites will offer individuals a better opportunity to prevent diseases and quickly evaluate the diagnosis and treatment processes when they are sick.

In our study, websites belonging to commercial and health portal typologies were the most common websites in all search results. Based on 10 Google searches, of the websites listed on the first page of the results three websites each were created by commercial, government and non-profit institutions. When compared based on their typologies, there was no significant difference between the top 10 websites and the remaining websites. However, there was a significant difference between website typologies and reliability scores. This significant difference is related to the high *JAMA* scores on websites from the health portal. There was no significant difference between the top 10 websites and the remaining websites in terms of reliability and quality. No significant difference was found between the readability indices of the information on the websites based on the website typologies. There was no significant difference in information readability indexes between the top 10 websites and the remaining websites, but a significant relationship was found between Alexa ranks. The websites that appear on the first page of Google search results got more visits, which naturally increased their Alexa ranks.

Websites created by commercial sources accounted for the largest number of websites. Consistent with our findings, there are studies on different subjects demonstrating a higher number of commercial websites in the literature.<sup>35-37</sup> It has been established that these sites only have financial goals, do not provide reliable information, and may mislead their users. In addition, we found that 3 of 10 websites on the first page of Google search results were created by commercial sources, consistent with the literature.<sup>38</sup> Internet users often visit the websites on the first page of Google results to access information. These websites, which are promoted by Google for financial reasons, may misinform their users. We believe that websites with reliable information that do not have any financial goals should be ranked higher by search engines.

We found the presence of HONcode in 32 (33.7%) of 95 websites. Arif *et al.*<sup>37</sup> found HONcodes in 17.9% of the websites in their study, whereas Grewal *et al.*<sup>39</sup> found HONcodes in 16% of the websites. Our findings are consistent with the literature in this aspect. We found a significant difference in HONcode presence in websites according to typology. This difference was caused by health portals. In the literature, Chumber *et al.*<sup>40</sup> similarly found that health portals contained more HONcodes. Many studies report that the presence of HONcode provides reliable and quality information.<sup>35,38</sup> In addition, we found websites with HONcode to have high DISCERN and *JAMA* scores. Accordingly, we can deduce that healthcare professionals can inform their patients to refer to websites with HONcodes when searching for information about cLBP on the internet.

We found the mean DISCERN score of the websites was found to be 'poor' at 40.65 (19.94). Similarly, Guo *et al.*<sup>9</sup> examined internet-based materials in the field of failed back spinal surgery and also found a mean DISCERN score of 35.26 (11.45). Moreover, there are studies in the literature reporting higher DISCERN scores.<sup>41,42</sup> However, the use of academic websites or websites of scientific journals in these studies led to high DISCERN scores as well as high *JAMA* scores and difficult

readability. Patients prefer more readable sources with less medical terminology when they try to obtain health-related information from the internet. In this sense, academic resources are tailored for use among health professionals and to contribute to science, rather than being used by patients.

We found no significant difference between the top 10 websites and other websites in terms of readability of information. Similar to our study, Bagcier *et al.*<sup>38</sup> and Kocyigit *et al.*<sup>43</sup> also found no significant difference between the two groups of websites in terms of readability of information. Higher readability of the information in the top 10 websites that are visited more often will help users to understand the information easily and quickly.

When the readability of all website typologies was compared, no statistically significant difference was found. The average readability results obtained in our study were found to be well above the 6th grade reading level recommended by the National Institute of Health.<sup>41</sup> In another study, Hendrick *et al.*<sup>11</sup> excluded academic websites from their analyses and found that the websites reporting information on aLBP had a moderate readability level than that understandable by the public. However, studies involving academic websites reported worse readability results.<sup>38,43</sup> We would like to point out that with easier readability levels, the contents of a website can reach wider audiences, and the power of information can be presented more effectively by providing information with a readability level that matches that of the public.

We found no significant relationship between popularity (Alexa) and readability indexes ( $p > 0.05$ ), however, a significant difference was found between Alexa and website typologies ( $p = 0.007$ ). This difference is related to Alexa values on non-profit sites. These findings are similar to those in the literature.<sup>42,43</sup> Non-profit institutions try to provide quality information to users without financial concerns and try to reach more users. We can explain the reasons for high visits by the fact that these websites have won the admiration of the audience. We found a significant difference when the top 10 websites were compared with other websites in terms of their Alexa values. Naturally, this significant difference can be explained by the fact that the top 10 websites are visited more often.

When the content of the websites was analyzed, it was determined that there were 77 (81.1%) websites with treatment-based content, followed by 69 (72.6%) websites with aetiology-based content. A statistically significant difference was found between website typologies and topics. Websites prepared by non-profit institutions and commercial websites mostly covered topics related to diagnosis, whereas websites prepared by commercial and professional institutions mostly covered topics related to symptoms. Bagcýer *et al.*<sup>38</sup> also reported that websites on myofascial pain mostly referred to treatment-related issues. Furthermore, physiotherapy-based content, including exercise programmes, was the most common treatment topic on these websites. Based on these results, it can be said that popular topics relevant to a particular disease find more space on websites.

Our study has limitations. We only searched for websites in English, used a single search engine, used only 'cLBP' as a search keyword, and only detected internet sites that used a single country's data network. Although there is no consensus on which index is ideal for assessing the readability of internet-based PEMs, the indices we used are among the most widely used. The websites were targeting an education level

considerably over the appropriate level, according to all of the metrics we analyzed.

We believe that when creating health-related websites for the public about cLBP, which is an important cause of disability, the language of the website should be checked according to the relevant readability indexes, the content on the website should have a readability level suitable to the average education level of the country or countries for which the information is intended, and the website should contain high quality and reliable information.

## REFERENCES

- Hayden JA, Ellis J, Ogilvie R, Malmivaara A, van Tulder MW. Exercise therapy for chronic low back pain. *Cochrane Database Syst Rev* 2021;**9**:CD009790.
- GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018;**392**:1789–858. Erratum in: *Lancet* 2019;**393**:e44.
- Grotle M, Brox JI, Veierød MB, Glomsrød B, Lønn JH, Vøllestad NK. Clinical course and prognostic factors in acute low back pain: Patients consulting primary care for the first time. *Spine (Phila Pa 1976)* 2005;**30**:976–82.
- Henschke N, Maher CG, Refshauge KM, Herbert RD, Cumming RG, Bleasel J, et al. Characteristics of patients with acute low back pain presenting to primary care in Australia. *Clin J Pain* 2009;**25**:5–11.
- Itz CJ, Geurts JW, van Kleef M, Nelemans P. Clinical course of non-specific low back pain: A systematic review of prospective cohort studies set in primary care. *Eur J Pain* 2013;**17**:5–15.
- Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. *Lancet* 2018;**391**:2356–67.
- Wenger HC, Cifu AS. Treatment of low back pain. *JAMA* 2017;**318**:743–4.
- Wang SW, Capo JT, Orillaza N. Readability and comprehensibility of patient education material in hand-related web sites. *J Hand Surg Am* 2009;**34**:1308–15.
- Guo WJ, Wang WK, Xu D, Qiao Z, Shi YL, Luo P. Evaluating the quality, content, and readability of online resources for failed back spinal surgery. *Spine (Phila Pa 1976)* 2019;**44**:494–502.
- Han A, Carayannopoulos AG. Readability of patient education materials in physical medicine and rehabilitation (PM&R): A comparative cross-sectional study. *PMR* 2020;**12**:368–73.
- Hendrick PA, Ahmed OH, Bankier SS, Chan TJ, Crawford SA, Ryder CR, et al. Acute low back pain information online: An evaluation of quality, content accuracy and readability of related websites. *Man Ther* 2012;**17**:318–24.
- Járomi M, Szilágyi B, Velényi A, Leidecker E, Raposa BL, Hock M, et al. Assessment of health-related quality of life and patient's knowledge in chronic non-specific low back pain. *BMC Public Health* 2021;**21** (Suppl 1):1479.
- Worldwide desktop market share of leading search engines from January 2010 to January 2022. Available at [www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/](http://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/) (accessed 1 Mar 2022).
- Basavakumar D, Flegg M, Eccles J, Ghezzi P. Accuracy, completeness and accessibility of online information on fibromyalgia. *Rheumatol Int* 2019;**39**:735–42.
- Arif N, Ghezzi P. Quality of online information on breast cancer treatment options. *Breast* 2018;**37**:6–12.
- Eysenbach G, Köhler C. 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. *BMJ* 2002;**324**:573–7.
- Zeldman J. Taking your talent to the web: A guide for the transitioning designer. Indianapolis:New Riders; 2001:1–449.
- Charnock D, Shepperd S, Needham G, Gann R. DISCERN: An instrument for judging the quality of written consumer health information on treatment choices. *J Epidemiol Community Health* 1999;**53**:105–11.
- Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor—Let the reader and viewer beware. *JAMA* 1997;**277**:1244–5.
- Weil AG, Bojanowski MW, Jamart J, Gustin T, Lévêque M. Evaluation of the quality of information on the internet available to patients undergoing cervical spine surgery. *World Neurosurg* 2014;**82**:e31–e39.
- Boyer C, Selby M, Appel RD. The health on the net code of conduct for medical and health web sites. *Stud Health Technol Inform* 1998;**52 Pt 2**:1163–6.
- Boyer C, Baujard V, Geissbuhler A. Evolution of health web certification through the HONcode experience. *Stud Health Technol Inform* 2011;**169**:53–7.
- Walsh T, Volsko T. Readability assessment of internet-based consumer health information. *Respir Care* 2008;**53**:1310–15.
- Garfinkle R, Wong-Chong N, Petrucci A, Sylla P, Wexner SD, Bhatnagar S, et al. Assessing the readability, quality and accuracy of online health information for patients with low anterior resection syndrome following surgery for rectal cancer. *Colorectal Dis* 2019;**21**:523–31.
- Calo WA, Gilkey MB, Malo TL, Robichaud M, Brewer NT. A content analysis of HPV vaccination messages available online. *Vaccine* 2018;**36**:7525–9.
- Sheats MK, Royal K, Kedrowicz A. Using readability software to enhance the health literacy of equine veterinary clients: An analysis of 17 American Association of Equine Practitioners' newsletter and website articles. *Equine Vet J* 2019;**51**:552–5.
- Huang G, Fang CH, Agarwal N, Bhagat N, Eloy JA, Langer PD. Assessment of online patient education materials from major ophthalmologic associations. *JAMA Ophthalmol* 2015;**133**:449–54.
- Yılmaz FH, Tutar MS, Arslan D, Çeri A. Readability, understandability, and quality of retinopathy of prematurity information on the web. *Birth Defects Res* 2021;**113**:901–10.
- Wald HS, Dube CE, Anthony DC. Untangling the web—the impact of Internet use on health care and the physician–patient relationship. *Patient Educ Couns* 2007;**68**:218–24.
- Bujnowska-Fedak MM, Węgierek P. The impact of online health information on patient health behaviours and making decisions concerning health. *Int J Environ Res Public Health* 2020;**17**:880.
- Paterick TE, Patel N, Tajik AJ, Chandrasekaran K. Improving health outcomes through patient education and partnerships with patients. *Proc (Bayl Univ Med Cent)* 2017;**30**:112–13.
- Agarwal N, Hansberry DR, Sabourin V, Tomei KL, Prestigiacomo CJ. A comparative analysis of the quality of patient education materials from medical specialties. *JAMA Intern Med* 2013;**173**:1257–9.
- Daraz L, Morrow AS, Ponce OJ, Farah W, Katabi A, Majzoub A, et al. Readability of Online Health Information: A meta-narrative systematic review. *Am J Med Qual* 2018;**33**:487–92.
- Eysenbach G, Kohler Ch. What is the prevalence of health-related searches on the world wide web? Qualitative and quantitative analysis of search engine queries on the internet. *AMIA Annu Symp Proc* 2003;**2003**:225–9.
- Yurdakul OV, Kilicoglu MS, Bagcier F. Evaluating the reliability and readability of online information on osteoporosis. *Arch Endocrinol Metab* 2021;**65**:85–92.
- Zhang D, Schumacher C, Harris MB, Bono CM. The quality and readability of information available on the internet regarding lumbar fusion. *Global Spine J* 2016;**6**:133–8.
- Arif N, Ghezzi P. Quality of online information on breast cancer treatment options. *Breast* 2018;**37**:6–12.
- Bagcier F, Yurdakul OV, Temel MH. Quality and readability of online information on myofascial pain syndrome. *J Body Mov Ther* 2021;**25**:61–6.
- Grewal P, Alagaratnam S. The quality and readability of colorectal cancer information on the internet. *Int J Surg* 2013;**11**:410–13.
- Chumber S, Huber J, Ghezzi P. A methodology to analyze the quality of health information on the internet: The example of diabetic neuropathy. *Diabetes Educ* 2015;**41**:95–105.
- Willen RD, Pipitone O, Daudfar S, Jones JD. Comparing quality and readability of online English language information to patient use and perspectives for common rheumatologic conditions. *Rheumatol Int* 2020;**40**:2097–103.
- Ng MK, Mont MA, Piuze NS. Analysis of readability, quality, and content of online information available for “stem cell” injections for knee osteoarthritis. *J Arthroplasty* 2020;**35**:647–51.e2.
- Kocuyigit BF, Koca TT, Akaltun MS. Quality and readability of online information on ankylosing spondylitis. *Clin Rheumatol* 2019;**38**:3269–74.