



DOI: 10.4274/qrheumatol.galenos.2025.07379

Rheumatol Q 2025;3(1):12-9

# QUALITY AND READABILITY OF CHATGPT'S RESPONSES TO THE MOST FREQUENTLY SEARCHED WORDS ABOUT FIBROMYALGIA ON GOOGLE TRENDS

Alper Uysal<sup>1</sup>, Ertürk Güntürk<sup>2</sup><sup>1</sup>Mersin City Training and Research Hospital, Clinic of Physical Medicine and Rehabilitation, Mersin, Türkiye<sup>2</sup>Adana City Training and Research Hospital, Clinic of Physical Medicine and Rehabilitation, Adana, Türkiye

## Abstract

**Aim:** We aimed to evaluate the quality and readability of ChatGPT's answers to frequently asked questions about fibromyalgia (FM).

**Material and Methods:** The most frequently searched terms related to FM were identified using Google Trends and entered into the ChatGPT-4 model in order of their rankings. The responses were categorized using the Ensuring Quality Information for Patients (EQIP) tool and evaluated based on quality and readability. Quality and readability were assessed with EQIP, the Flesch-Kincaid Grade Level (FKGL), and the Flesch-Kincaid Reading Ease (FKRE).

**Results:** According to Google Trends data, the search frequency for the term "FM" increased from 2004 to 2024, with a peak of 100% in March 2020. ChatGPT's responses were assessed in terms of both quality and readability, revealing notable shortcomings. The average scores for EQIP, FKGL, and FKRE were 39.89, 13.29, and 21.41, respectively. Furthermore, the statistical analysis among the four categories showed no significant variations in EQIP, FKGL, and FKRE scores.

**Conclusion:** Our study revealed significant deficiencies in the quality of ChatGPT's responses regarding FM. There is a need for more understandable and reliable information to improve communication in healthcare.

**Keywords:** Fibromyalgia, Google trends, ChatGPT, EQIP tool, Flesch-Kincaid Grade Level, Flesch-Kincaid Reading Ease

## INTRODUCTION

Fibromyalgia (FM) is a relatively common chronic pain syndrome within the general population, with a global prevalence rate of approximately 2-3%. In addition to chronic widespread pain, the FM clinical profile is characterized by fatigue, sleep disturbances, cognitive impairment, autonomic dysfunction, somatic symptoms, and psychiatric disorders, all occurring without any

underlying serious medical conditions (1). The pathogenesis of FM is incompletely understood; however, it is thought to involve a combination of genetic predisposition, stressful environmental factors, inflammatory processes, and central mechanisms such as pain centralization (2). The diagnosis of FM is made primarily based on clinical findings. Although physical examination and laboratory tests may not be definitive for diagnosis, they are

**Address for Correspondence:** Alper Uysal, Mersin City Training and Research Hospital, Clinic of Physical Medicine and Rehabilitation, Mersin, Türkiye

**E-mail:** alperuysal82@gmail.com **ORCID ID:** orcid.org/0000-0002-4114-1649

**Received:** 17.11.2024 **Accepted:** 06.01.2025 **Publication Date:** 21.03.2025

**Cite this article as:** Uysal A, Güntürk E. Quality and readability of ChatGPT's responses to the most frequently searched words about fibromyalgia on google trends. Rheumatol Q. 2025;3(1):12-9



crucial in excluding other conditions. The diagnosis is ultimately established through a detailed medical history (1,2). Patient education, physical exercise, pharmacological agents, and cognitive behavioral therapy are the four main components of the treatment (3).

The Internet has become an essential and irreplaceable resource for health-related information. It is now a significant source of information for patients with FM. The demand for information is heightened by the absence of specific diagnostic tools, evidence-based treatment recommendations for FM, and the ongoing debates surrounding the condition. Barriers to understanding specialized medical terminology and the chronic nature of FM further contribute to this increased demand (4). Research indicates that individuals living with FM frequently turn to the Internet to learn more about the disease and to support shared decision-making processes with their healthcare providers (4-6).

ChatGPT, developed by OpenAI (United States), is a highly advanced language model that leverages deep learning techniques to generate responses that closely mimic human language patterns. It is adept at understanding the subtleties and complexities of human communication, enabling it to produce contextually accurate and relevant responses across a wide range of topics (7).

In summary, FM exemplifies the challenges of patient education due to the lack of definitive diagnostic tests, the absence of universally accepted treatment guidelines, and widespread misconceptions about the condition. These factors often lead patients to seek information online, where the quality and accuracy of content vary greatly, and this may pose the risk of misinformation. FM was chosen as the focus disease for this study because it represents an especially relevant condition to evaluate the potential and limitations of AI-driven tools like ChatGPT in addressing global health information needs.

In this study, we aimed to evaluate the quality and readability of ChatGPT's answers to frequently asked questions about FM.

## MATERIAL AND METHODS

This study, conducted on August 24, 2024, at the PM&R clinics of Mersin City Training and Research Hospital and Adana City Training and Research Hospital, was exempt from ethics committee approval as it only utilized online information and did not involve any human or animal participants. Consequently, no informed consent form was required. In compliance with the ethical principles of the Declaration of Helsinki, this study adheres to all required ethical standards. As no human or animal data were utilized, obtaining approval from an ethics committee was not necessary (8,9).

In this study, English was used as the primary language for conducting Google Trends analyses with the term "FM" generating ChatGPT responses, and evaluating these responses, using the Flesch-Kincaid Reading Ease (FKRE) and Flesch-Kincaid Grade Level (FKGL). This choice was made because English is a widely used international language, particularly in the fields of health, medicine, and scientific communication. By utilizing English terms, the study aimed to capture search behaviors from a broader audience across various regions and countries.

Before searches were made, all personal web browsing history was preemptively cleared to prevent interaction. The Google Trends tool was utilized to identify the top searched terms related to FM. Global searches under the health category, spanning from 2004 to August 25, 2024, were selected to determine top keywords associated with FM. The "most relevant" option was selected from the related questions section, while the "subregion" option was chosen from the geographical areas of interest section.

The top 25 searched terms were noted, encompassing a broad spectrum of topics in Google's online searches. Two terms "ms" and "me" identified during the process were excluded from the analysis as they were deemed irrelevant to the main context of FM-related searches.

The term "ICD-10 FM" was included as it appeared among the most frequently searched keywords on Google Trends. Although ICD-10 codes are primarily utilised by healthcare professionals, patients may also encounter these codes in their medical reports and seek to understand their meaning through online searches. Including this keyword ensured a comprehensive analysis of search behaviours and helped minimise selection bias, accurately reflecting the diversity of queries related to FM.

The identified keywords were entered into the ChatGPT-4 model sequentially, according to their rankings in Google Trends, and are presented in Table 1 in the same order.

Before evaluating ChatGPT's performance with the keywords, the web browsing history was cleared, and separate chat pages were opened for each keyword. This approach helped to prevent potential interactions. The responses generated by ChatGPT-4 were systematically catalogued to evaluate the quality of information, readability, and comprehensiveness.

The 23 keywords were divided into four categories according to the criteria specified in the "Ensuring Quality Information for Patients" tool (EQIP): condition or illness; drug, medication, or product; treatment or management; and diagnosis, testing, or procedures (Table 1). The EQIP tool takes into account various parameters, including clarity of information, writing quality, accuracy, reliability, and comprehensiveness. Each of the 20

**Table 1. Keywords and the category of the topic based on EQIP**

Rank	Keyword	Category of the topic based on EQIP
1	What are the symptoms of fibromyalgia	Condition or illness
2	Is fibromyalgia real	Condition or illness
3	What causes fibromyalgia	Condition or illness
4	Symptoms of fibromyalgia in women	Condition or illness
5	How is fibromyalgia diagnosed	Diagnosis, testing, or procedures
6	How to treat fibromyalgia	Treatment or management
7	Amitriptyline for fibromyalgia	Drug, medication, or product
8	Is fibromyalgia a disability	Condition or illness
9	Cymbalta for fibromyalgia	Drug, medication, or product
10	Fibromyalgia meds	Drug, medication, or product
11	Lyrica fibromyalgia	Drug, medication, or product
12	Define fibromyalgia	Condition or illness
13	Fibromyalgia in men	Condition or illness
14	Gabapentin for fibromyalgia	Drug, medication, or product
15	Lyrica for fibromyalgia	Drug, medication, or product
16	Naltrexone for fibromyalgia	Drug, medication, or product
17	Savella for fibromyalgia	Drug, medication, or product
18	Fibromyalgia flare	Condition or illness
19	Is fibromyalgia autoimmune	Condition or illness
20	Fibromyalgia and chest pain	Condition or illness
21	Is fibromyalgia hereditary	Condition or illness
22	ICD 10 fibromyalgia	Condition or illness
23	Allodynia	Condition or illness

EQIP: Ensuring Quality Information for Patients, ICD: International classification of diseases

parameters in the EQIP scale was assigned a score of 1 for a “yes” response, 0.5 for a “partial” response, and 0 for a “no” or “not applicable (N/A)” response. The scores for each parameter were then summed and divided by the total number of parameters. Finally, the resulting EQIP score was calculated and expressed as a percentage. According to EQIP scores, texts can be classified as well-written and high quality, (76% to 100%), good quality with minor problems, (51% to 75%), having serious quality issues, (26% to 50%), or having severe quality issues, (0% to 25%) (9,10). To minimize bias in the calculation of EQIP, the recorded responses were independently evaluated by two physical medicine and rehabilitation specialists in separate settings. In cases of discrepancy, the average values were taken.

The readability of the texts was analyzed using two key metrics: the FKGL and the FKRE. These parameters were employed to assess the complexity and accessibility of the text content.

The FKGL was determined through a series of specific calculations. Initially, the total number of words was divided by the total number of sentences, and this value was multiplied by 0.39. Subsequently, the total number of syllables was divided by the total number of words, with the resulting figure multiplied by 11.8. The results from these calculations were then added together, and finally, 15.59 was subtracted from the sum to yield the FKGL score. A lower grade level score signifies that the text is easier to understand, whereas a higher score indicates greater linguistic complexity and requires a more advanced level of comprehension.

The FKRE formula assesses a document’s readability by first calculating the average sentence length, which is then multiplied by 1,015, and the average number of syllables per word, which is multiplied by 84.6. The difference between these two products is subtracted from 206,835 to yield the document’s reading ease score. Higher scores reflect text that is easier to read, whereas a

score of 30 or below indicates that the content is likely suitable for a reading level appropriate for a college graduate (11).

### Statistical Analysis

The statistical program used in the study was IBM SPSS Statistics for Windows, Version 22.0 (Armonk, NY, USA). The normality of continuous variables was assessed using the Kolmogorov-Smirnov test and histogram plots. Categorical variables were presented as numbers (n) and percentages (%). For continuous variables that followed a normal distribution, Independent Samples t-test was used for comparisons between two groups, while One-Way Analysis of Variance was employed for comparisons among four groups. For continuous variables that did not follow a normal distribution, comparisons among four groups were performed using the Kruskal-Wallis test. In cases where the Kruskal-Wallis test identified significant differences, pairwise comparisons were conducted using the Mann-Whitney U test. Groups with insufficient sample sizes (n=1) were excluded from the analysis. Bonferroni correction was not applied for pairwise comparisons, as only a single Mann-Whitney U test was conducted. Intrarater reliability for each researcher and interrater reliability between the researchers were analysed using Cohen's kappa coefficient for categorical data. A significance level of  $p < 0.05$  was considered statistically significant for all analyses.

## RESULTS

The graph illustrates the temporal variation in search frequency for the term "FM" according to Google Trends data from 2004 to 2024 (Figure 1). In January 2004, the relative search interest for "FM" stood at 56%, while by August 2024, this interest had risen to 73%. Notably, the graph features a pronounced spike in March 2020, when the relative search interest reached 100%.

As depicted in the graph, Norway, Puerto Rico, and the United Kingdom are the top three countries where the term "FM" has

been most frequently searched. Geographically, the highest search interest is concentrated in North America, Western and Northern European countries, and Australia (Figure 2).

Table 1 presents the 23 keywords obtained from Google Trends along with their corresponding categories as determined by the EQIP analysis.

The minimum, maximum, mean, and standard deviation values for EQIP, FKGL, and FKRE, as well as the percentages for the category of the topic based on EQIP, are presented in Table 2. Based on the EQIP analysis, the most frequently observed categories for this topic are "condition or illness," which accounts for 57%, and "drug, medication, or product," which represents 35% (Figure 3).

No statistically significant differences were found between Group 1 and Group 2 in terms of EQIP, FKGL, and FKRE values. In Groups 3 and 4, the standard deviation could not be calculated due to the response count being only one. Additionally, the statistical analysis across the four groups revealed no significant differences in EQIP, FKGL, and FKRE values (Table 3). The obtained kappa value (0.79) represents a high level of interrater reliability ( $p < 0.001$ ). The intrarater reliability analysis conducted after 109 days yielded kappa values of 0.87 for researcher 1 (Alper Uysal) and 0.85 for researcher 2 (Ertürk Güntürk), both statistically significant ( $p < 0.001$ ).

## DISCUSSION

Our study revealed significant issues in the quality of responses provided by ChatGPT. The FKGL corresponded to a 13-year education level, while the FKRE score indicated that the texts were at a difficult readability level.

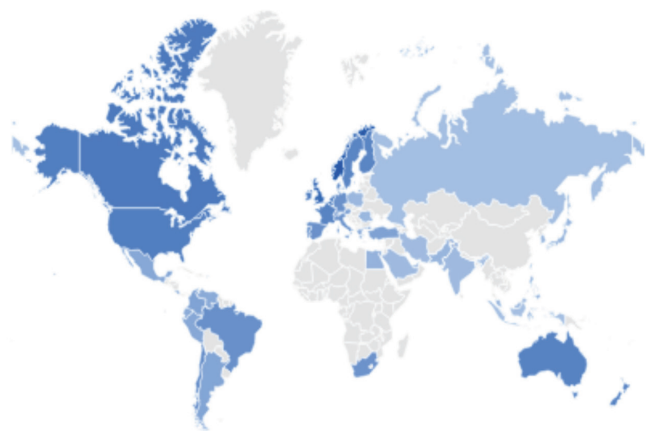
FM significantly impacts individuals' lives, often manifesting as chronic widespread pain, sleep disturbances or non-restorative sleep, physical exhaustion, and cognitive difficulties.



**Figure 1.** Temporal change in search frequency for the term "fibromyalgia"

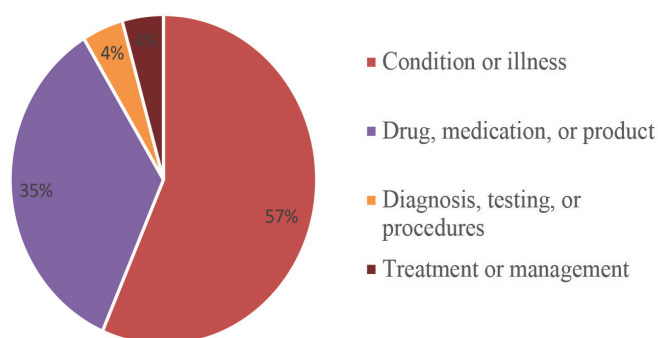
This condition involves a range of somatic and psychological symptoms and is frequently accompanied by co-morbid illnesses such as functional somatic syndromes (for instance, irritable bowel syndrome), anxiety and depressive disorders, and rheumatic diseases. These symptoms and associated conditions can profoundly complicate daily living, substantially diminishing quality of life (2).

When analysing Google Trends data for the term “FM” from 2004 to 2024, a significant spike is observed in March 2020, with the relative search interest peaked at 100%. On 11 March 2020, the World Health Organization declared COVID-19 a global pandemic (12), and by 13 March 2020, Europe was reported as the new epicentre of the crisis (13). The spike in searches for FM during this period may be linked to the heightened anxiety,



**Figure 2.** Global search interest in fibromyalgia: 2004-2023

**Category of the Topic Based on EQIP**



**Figure 3.** Category of the topic based on EQIP  
EQIP: Ensuring Quality Information for Patients

**Table 2. Quantitative properties of parameters**

Parameters	Minimum	Maximum	Mean	Standart deviation	n
EQIP	25.00	56.25	39.89	6.20	23
FKGL	11.10	17.50	13.29	1.56	23
FKRE	5.84	39.90	21.41	9.37	23
<b>Percentage</b>					
<b>Category of the topic based on EQIP</b>	Condition or illness			57%	13
	Drug, medication, or product			35%	8
	Diagnosis, testing, or procedures			4%	1
	Treatment or management			4%	1

n: Number, EQIP: Ensuring Quality Information for Patients, FKGL: Flesch-Kincaid Grade Level, FKRE: Flesch-Kincaid Reading Ease

**Table 3. Comparison of parameters between groups**

	Group 1 (n=13)	Group 2 (n=8)	p*	Group 3 (n=1)	Group 4 (n=1)	p**
	Mean ± SD	Mean ± SD		Mean	Mean	
FKGL	13.62±1.72	12.86±1.29	0.298	11.60	14.30	0.457
FKRE	19.26±8.27	24.74±9.94	0.188	33.60	10.60	0.195
	Median (min.-max.)	Median (min.-max.)		Median	Median	
EQIP	37.50 (25.00-56.25)	44.38 (32.50-45.13)	0.059	34.38	43.75	0.136

\*p-value for Group 1 and Group 2, \*\*p-value for all groups, n: Number, SD: Standart deviation, EQIP: Ensuring Quality Information for Patients, FKGL: Flesch-Kincaid Grade Level, FKRE: Flesch-Kincaid Reading Ease, Min.-max.: Minimum-maximum

depression, and stress caused by the pandemic, which could have exacerbated FM symptoms. Notably, the surge in FM searches coincided with the rapid increase in COVID-19 cases, suggesting a possible correlation between the pandemic's impact on mental health and the worsening of FM symptoms. When evaluating this relationship, it is also important to consider Long COVID, which presents symptoms that closely resemble those of FM (14,15).

Despite such a complicated clinical picture, no cure currently exists for FM, so treatment is centred on improving the patient's ability to function while managing pain and other associated symptoms (16). Moreover, given the complicated aspects of FM, numerous myths and misconceptions abound. All of these leads patients to actively seek reliable information (17). Patients may wish to gain a deeper understanding of their condition and the available treatment options to actively engage in the decision-making process. An informed patient is better equipped to contribute meaningfully to decisions about their care and may experience reduced anxiety as a result. Conversely, without access to quality information, patients may struggle to discuss their findings coherently with their doctors and may be unable to make well-informed decisions (18,19). 72% of internet users seek medical information online (20). However, it is argued that the quality of web information is often poor and is presented in a way that increases the likelihood of misunderstanding (4,21). Basavakumar et al. (21) have shown that online resources concerning FM, including its etiology, symptoms, comorbidities, and management, are frequently incomplete, with content that may be difficult to access and susceptible to misinterpretation. Ozsoy-Unubol and Alanbay-Yagci (19) investigated the quality of online information on YouTube concerning FM. They found that, despite its variability, the content on the YouTube platform was generally of poor quality. Moreover, they suggested that this poor-quality information could mislead patients and potentially harm the doctor-patient relationship.

Previous research has highlighted that online resources for FM, including those on platforms like YouTube, are often incomplete or of poor quality (19,21). Such deficiencies can mislead patients and hinder effective communication with healthcare providers. Although ChatGPT provides information that is quickly and easily accessible, our findings suggest that the quality and clarity of its responses are insufficient to fully meet patient needs. For example, EQIP scores for ChatGPT responses were lower than the threshold for high-quality information.

The integration of ChatGPT into our daily experiences has led to its widespread adoption across various fields, including medicine. Beyond its use in medical education, research, and clinical practice, ChatGPT has also been employed in patient education and information (22).

In this context, studies have been conducted on various diseases, including osteoarthritis, cerebral palsy (CP), and osteoporosis (8,9,23). Ata et al. (8) assessed the reliability and utility of ChatGPT's responses concerning cerebral palsy and determined that it serves as a reliable, though partially useful, source of information. They emphasized the importance of patients and their families verifying the medical information obtained by consulting their healthcare providers. Erden et al. (9) assessed the quality, readability, and comprehensibility of the information provided by ChatGPT concerning osteoporosis. Their findings indicated that the quality and readability of ChatGPT's information were insufficient for effective health management. Yang et al. (23) determined that ChatGPT's responses do not always fully align with recommendations from evidence-based clinical practice guidelines, and they suggested that both healthcare professionals and patients should approach the guidance offered by AI platforms with measured expectations. They should recognize the current limitations in providing clinically sound advice. Consistent with existing literature on other diseases (8,9,23), our study revealed that the quality of responses provided by ChatGPT was insufficient. Additionally, we determined that the understandability of the content was categorized as difficult. We believe that the optimal approach to information transfer is to convey highly reliable information in a manner that is comprehensible to all segments of society.

In diseases like FM, where the etiopathogenesis, diagnosis, and treatment are complex, excessive reliance on medical information obtained through artificial intelligence and the internet could have adverse effects on patient health. While these tools offer quick and easy access to information, the principle that treatment should be tailored to the individual rather than the disease itself implies that general information may be insufficient or even inaccurate in complex and complicated cases (9,24).

This study aimed to evaluate ChatGPT's ability to generate globally relevant health-related responses by focusing on English-language keywords and international search trends. While regional trends are valuable, the study prioritized global applicability, leaving localized analyses for future research.

Future studies should explore keywords in other languages to understand how language and cultural differences influence the quality and accessibility of AI-generated healthcare content. This could help create more inclusive AI tools, addressing disparities in reliable health information for non-English-speaking populations.

The main limitations of the study are the exclusive focus on English terms, which could constrain the findings, and the

selection of only 23 keywords. Expanding the keyword lists and incorporating other languages in future studies could yield more comprehensive results.

## CONCLUSION

Our study revealed significant deficiencies in the quality and clarity of FM-related responses generated by ChatGPT, and these deficiencies raise concerns about its suitability for effective health communication. Healthcare professionals should guide patients in interpreting AI-generated information and verifying it with evidence-based sources. AI developers must enhance readability and alignment with clinical guidelines, to improve reliability. Patients should critically evaluate online health content and consult healthcare providers for accurate advice. Refining algorithms is crucial to advancing the accuracy and accessibility of AI-generated responses for more effective health communication. Future research should examine non-English keywords to enhance AI tools' inclusivity and address health information disparities across different languages and cultures.

## Ethics

**Ethics Committee Approval:** As no human or animal data were utilized, obtaining approval from an ethics committee was not necessary.

**Informed Consent:** No informed consent form was required.

## Footnotes

### Authorship Contributions

Surgical and Medical Practices: A.U., E.G., Concept: A.U., E.G., Design: A.U., E.G., Data Collection or Processing: A.U., E.G., Analysis or Interpretation: A.U., E.G., Writing: A.U., E.G.

**Conflict of Interest:** The authors have no conflicts of interest to declare.

**Financial Disclosure:** The authors declared that this study received no financial support.

## REFERENCES

1. Sarzi-Puttini P, Giorgi V, Marotto D, Atzeni F. Fibromyalgia: an update on clinical characteristics, aetiopathogenesis and treatment. *Nat Rev Rheumatol*. 2020;16:645-60.
2. Häuser W, Ablin J, Fitzcharles MA, et al. Fibromyalgia. *Nat Rev Dis Primers*. 2015;1:15022.
3. Llädser AN, Montesó-Curto P, López C, et al. Multidisciplinary rehabilitation treatments for patients with fibromyalgia: a systematic review. *Eur J Phys Rehabil Med*. 2022;5:76-84.
4. Daraz L, Macdermid JC, Wilkins S, Gibson J, Shaw L. The quality of websites addressing fibromyalgia: an assessment of quality and readability using standardised tools. *BMJ Open*. 2011;1:e000152.
5. Crooks VA. "I Go on the internet; I always, You know, check to See What's New": chronically ill Women's use of online health information to shape and inform doctor-patient interactions in the space of care provision. *ACME*. 2006;5:50-69.
6. Daraz L. Information availability and needs of people living with fibromyalgia. *Diss Abstr Int B Sci Eng*. 2012:72.
7. Brown TB, Mann B, Ryder N, et al. Language Models are Few-Shot Learners. *arXiv preprint*. 2020;arXiv:2005.
8. Ata AM, Aras B, Yılmaz Taşdelen Ö, Çelik C, Çulha C. Evaluation of informative content on cerebral palsy in the era of artificial intelligence: the value of ChatGPT. *Phys Occup Ther Pediatr*. 2024;44:605-14.
9. Erden Y, Temel MH, Bağcıer F. Artificial intelligence insights into osteoporosis: assessing ChatGPT's information quality and readability. *Arch Osteoporos*. 2024;19:17.
10. Moulton B, Franck LS, Brady H. Ensuring quality information for patients: development and preliminary validation of a new instrument to improve the quality of written health care information. *Health Expect*. 2004;7:165-75.
11. Boles CD, Liu Y, November-Rider D. Readability levels of dental patient education brochures. *J Dent Hyg*. 2016;90:28-34.
12. WHO T. WHO Director-General's opening remarks at the media briefing on COVID-19-11 March 2020. Geneva; 2020.
13. Carroll W, Strenger V, Eber E, et al. European and United Kingdom COVID-19 pandemic experience: the same but different. *Paediatr Respir Rev*. 2020;35:50-6.
14. Fialho MFP, Brum ES, Oliveira SM. Could the fibromyalgia syndrome be triggered or enhanced by COVID-19? *Inflammopharmacology*. 2023;31:633-51.
15. Mariette X. Long COVID: a new word for naming fibromyalgia? *Ann Rheum Dis*. 2024;83:12-4.
16. Hoffman DL, Dukes EM. The health status burden of people with fibromyalgia: a review of studies that assessed health status with the SF-36 or the SF-12. *Int J Clin Pract*. 2008;62:115-26.
17. Bragazzi NL, Amital H, Adawi M, et al. What do people search online concerning the "elusive" fibromyalgia? Insights from a qualitative and quantitative analysis of Google Trends. *Clin Rheumatol*. 2017;36:1873-8.
18. Longtin Y, Sax H, Leape LL, Sheridan SE, Donaldson L, Pittet D. Patient participation: current knowledge and applicability to patient safety. *Mayo Clin Proc*. 2010;85:53-62.
19. Ozsoy-Unubol T, Alanbay-Yagci E. YouTube as a source of information on fibromyalgia. *Int J Rheum Dis*. 2021;24:197-202.
20. Fox S, Duggan M. Health online 2013. *Health*. 2013;2013:1-55.

21. Basavakumar D, Flegg M, Eccles J, Ghezzi P. Accuracy, completeness and accessibility of online information on fibromyalgia. *Rheumatol Int.* 2019;39:735-42.
22. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare (Basel).* 2023;11:887.
23. Yang J, Ardavanis KS, Slack KE, Fernando ND, Della Valle CJ, Hernandez NM. Chat Generative Pretrained Transformer (ChatGPT) and Bard: artificial intelligence does not yet provide clinically supported answers for hip and knee osteoarthritis. *J Arthroplasty.* 2024;39:1184-90.
24. Tinetti ME, Fried T. The end of the disease era. *Am J Med.* 2004;116:179-85.