Assessing the Quality, Readability, and Comprehensibility of Fibromyalgia and Treatment Information Provided by ChatGPT

Tugce Pasin[1] , Bilinc Dogruoz Karatekin[2] , Ozge Pasin[3]

ABSTRACT

This study aimed to investigate the quality, readability, and comprehensibility of responses provided by ChatGPT to the most frequently searched topics using the keywords "fibromyalgia" and "fibromyalgia treatment." Popular keywords for "fibromyalgia" and "fibromyalgia treatment" were obtained using Google Trends and entered into ChatGPT 3.5. The Ensuring Quality Information for Patients (EQIP), Flesch-Kincaid Reading Ease (FKRE), and Flesch-Kincaid Grade Level (FKGL) scores were calculated for the generated responses. These scores were assessed by two experts, and comparisons were made between the fibromyalgia and fibromyalgia treatment keywords. General fibromyalgia information had a mean FKRE score of -2.98 ± 13.53. The mean FKGL score was 19.18 ± 2.36. The mean EQIP score was 41.60 ± 9.18, reflecting high-quality information. Statistically significant differences were observed in EQIP scores among different categories, with the "Drug, Medication, or Product" group scoring lower than other groups. Comparisons between fibromyalgia and fibromyalgia treatment keywords showed that general fibromyalgia information had higher FKRE scores, while treatment-related content had higher FKGL scores, indicating greater complexity. The findings highlight that while general information about fibromyalgia provided by ChatGPT is more readable, texts concerning fibromyalgia treatment are more complex.

Keywords

ChatGPT; language model; ChatGPT 3.5; natural language processing; fibromyalgia

INTRODUCTION

Artificial intelligence (AI) and natural language processing (NLP) technologies have rapidly advanced in recent years and are being widely used as information sources in various fields [1-3].

Recently, ChatGPT have begun to be widely used as a source of information in many different fields. They have the potential to provide information and services to people, especially in many fields such as health, education, finance, technology, art and law [4]. In the field of healthcare, it is used for tasks such as assessing patients' symptoms, providing information about healthcare services, and providing support to healthcare professionals [5].

These models have various advantages and disadvantages [6]. Large language models like ChatGPT can provide information on various topics based on a vast knowledge base. They enable quick and easy access for users. AI-based systems are always available, allowing individuals to access information at any time. This provides a significant advantage

1. Tugce Pasin, Goztepe Prof Dr Suleyman Yalcin City Hospital, Physical Medicine and Rehabilitation, Istanbul, Turkey. E-mail: tugcepasin@hotmail.com
2. Bilinc Dogruoz Karatekin, Goztepe Prof Dr Suleyman Yalcin City Hospital, Physical Medicine and Rehabilitation, Istanbul, Turkey. Email: bilincdogruoz@hotmail.com
3. Ozge Pasin, Bezmialem Foundation University Faculty of Medicine, Department of Biostatistics, Istanbul, Turkey. Email: opasin@bezmialem.edu.tr.

Correspondence

Bilinc Dogruoz Karatekin, Goztepe Prof Dr Suleyman Yalcin City Hospital, Physical Medicine and Rehabilitation, Istanbul, Turkey. Email: bilincdogruoz@hotmail.com

in accessing urgent information. They can provide personalized content based on user preferences. This enhances user experience and facilitates more effective access to information. However, the accuracy of AI systems may be limited, especially in complex or highly specific topics. This can create reliability issues in areas requiring precision. Also AI-based systems may not fully replicate human interaction and may be inadequate in areas requiring personal experience. In healthcare topics, especially in cases requiring personalized diagnosis and treatment, AI-based systems may have limited competence. They cannot completely replace the expertise of doctors and specialists [7].

Artificial intelligence and systems like ChatGPT serve as significant information sources. However, they have shortcomings in terms of accuracy, personalization, and ethical considerations. Therefore, the use of these systems should be carefully evaluated, and human expertise and experience should be sought when necessary [8].

Fibromyalgia, a frequently encountered condition that lacks definitive diagnostic tests, often prompts fibromyalgia patients to search for information about the disease online [9]. Fibromyalgia is characterized by widespread body pain, fatigue, insomnia, emotional stress, and other symptoms. However, these symptoms can also be associated with other health issues, making it challenging to diagnose fibromyalgia conclusively. This uncertainty leads patients to seek information from various sources to understand and manage their symptoms [10]. The internet serves as a vital resource for fibromyalgia patients to research their symptoms, learn about treatment options, and connect with individuals who share similar experiences [9]. However, caution is advised regarding the accuracy and reliability of this information. Expert advice and information published by reputable health organizations play a crucial role in ensuring that patients access correct and trustworthy information.

Therefore this study aimed to investigate the quality, readability, and comprehensibility of the responses provided by ChatGPT to the most frequently searched topics using the keywords "fibromyalgia" and "fibromyalgia treatment."

MATERIALS AND METHODS

The preliminary search was conducted using the keywords "fibromyalgia" and "fibromyalgia treatment"

in Google Trends tool. Sixteen popular keywords were found for "fibromyalgia," and 22 popular keywords were found for "fibromyalgia treatment." These keywords were individually queried on the same day by 2 researchers who are both experts in this topic using the free version of ChatGPT 3.5. The responses provided by ChatGPT 3.5 were recorded and evaluated using The Flesch-Kincaid Reading Ease Score (FKRE), The Flesch-Kincaid Grade Level Score (FKGL), and The Ensuring Quality Information for Patients (EQIP).

The Flesch-Kincaid Reading Ease Score (FKRE) is a tool commonly used to assess the readability of written texts, including health-related materials. It provides a numerical score based on factors such as sentence length and average number of syllables per word, with higher scores indicating easier readability. The validity of FKRE lies in its ability to objectively quantify the ease with which a text can be read and understood by the target audience. This score is particularly valuable in healthcare contexts, where clear communication of complex medical information is crucial for patient comprehension and decision-making. Researchers and healthcare professionals often rely on FKRE to ensure that health materials are presented in a manner that is accessible and understandable to patients and the general public [11].

The Flesch-Kincaid Grade Level Score (FKGL) is a readability metric commonly used to assess the complexity of written texts, including those in the healthcare domain. It calculates the reading level required to understand a piece of text, with lower scores indicating simpler language suitable for lower-grade readers. The validity of FKGL lies in its objective measurement of text complexity, which can help writers tailor their content to specific audience demographics or literacy levels. In healthcare, FKGL is valuable for ensuring that medical information is presented in a manner that is comprehensible to patients with varying levels of health literacy. By using FKGL, healthcare professionals and researchers can enhance the accessibility and effectiveness of health communication materials, ultimately improving patient understanding and decision-making [11].

The Ensuring Quality Information for Patients (EQIP) questionnaire is a scale used to assess the quality of health information for patients. Its validity is utilized to measure the accuracy, reliability, and comprehensibility of health information. EQIP ensures that health

information is easily understandable by patients and is based on reliable and high-quality sources. Therefore, the EQIP questionnaire is considered an effective tool in evaluating the quality of health information sources. Researchers aim to enhance the quality of healthcare services by using the EQIP questionnaire to measure the accuracy and comprehensibility of information accessed by patients [12].

Statistical Analysis

Descriptive statistics for quantitative variables including mean, standard deviation, median, minimum, and maximum were provided in the study. The normality of quantitative variables was assessed using the Shapiro-Wilk test. Variance homogeneity was evaluated using the Levene test. The relationships between quantitative variables were examined using Pearson correlation analysis. Mean comparisons between two groups were evaluated using the independent samples t-test (Student's t-test). The Kruskal-Wallis test was used for mean comparisons among more than two groups. The Dunn test was utilized as a post hoc test for multiple comparisons. A statistical significance level of 0.05 was adopted, and the calculations were performed using the SPSS software package (Version 26.0, IBM Corp., Armonk, NY).

RESULTS

The success of artificial intelligence in this area was evaluated by writing the popular keywords for "fibromyalgia" and "fibromyalgia treatment" topics to ChatGPT 3.5, Google Trends was used to obtain the popular keywords and the obtained popular keywords are presented in Table 1. Additionally, Table 1 provides the category classifications of the EQIP questionnaire. For fibromyalgia, the most searched countries on Google Trends were United Kingdom, Lebanon, Australia, Ireland, and New Zealand in order. For fibromyalgia treatment, the observed countries were Australia, United Kingdom, United States, and India. The EQIP scores were evaluated by two experts in the field. It was observed that there was a high level of agreement between the scores obtained from the researchers' EQIP evaluations (r=0.918; p<0.001).

Table 2 includes EQIP, FKRE, and FKGL scores obtained for definitions and explanations derived from ChatGPT 3.5 results. These values are typically used in statistical analyses and to measure the readability of texts. A high Flesch-Kincaid Readability Score represents an easily understandable text, while a low score indicates a more difficult-to-understand text. The Ensuring Quality Information for Patients (EQIP) score reflects the quality of health information, with higher values representing more reliable and high-quality information.

Descriptive statistics of the scores, the mean, median, standard deviation, minimum, and maximum were provided. When examining the FKRE score, it was observed that the mean score was -2.98 ± 13.530. This score represents the Flesch-Kincaid Readability Score used to measure the readability of a text. It generally ranges from 0 to 100, with higher values indicating easier readability. The fact that both the mean and median values are negative in this table indicates that these texts have a significantly difficult readability level. The high standard deviation also indicates that the scores are widely distributed.

The mean FKGL score was observed to be 19.18 ± 2.363. This score represents the Flesch-Kincaid Grade Level Score used to measure the comprehensibility of a text. As a result, these Flesch-Kincaid Reading Level statistics indicate that an academic article is generally suitable for an undergraduate-level readership, but certain sections may require simpler or more complex language. It can be concluded that the article is generally not technical and not difficult to understand, but some parts may require more expertise.

The mean EQIP score was obtained as 41.60 ± 9.182. This score is used to measure the level of providing quality information to patients. The values in this table generally reflect a high level of quality. Since the average and median values are high, it is understood that this information is generally of high quality. The standard deviation is at a moderate level, indicating that the scores are somewhat dispersed but mostly high.

Table 3 presents descriptive statistics and p-values obtained from comparing groups in terms of EQIP, FKRE, and FKGL scores among EQIP categories. According to Table 3, no statistically significant difference was observed between groups in terms of FKRE and FKGL scores among EQIP categories (p=0.300; p=0.108). However, a statistically significant

difference was observed among EQIP categories in terms of EQIP scores (p=0.004). The EQIP scores of questions in the Drug, Medication, or Product group was found to be significantly lower than the scores of the Discharge or Aftercare, Condition or Illness, and Test, Operation, Investigation, or Procedure groups (p=0.001; 0.014; p=0.030, respectively). The EQIP scores of the Discharge or Aftercare group was found to be significantly higher than the Miscellaneous group (p=0.010). No significant differences were observed among other groups in terms of EQIP scores (all p>0.05).

Table 4 compares the EQIP, FKRE, and FKGL scores obtained for Fibromyalgia and Fibromyalgia Treatment Keywords. Descriptive statistics and p-values resulting from the comparison are provided. According to Table 4, no statistically significant difference was observed between the two groups in terms of EQIP scores (p=0.147). It was observed that the FKRE scores in the Fibromyalgia group was statistically significantly higher than that in the Fibromyalgia Treatment group (p=0.002). When evaluated in terms of FKGL, the scores of the Fibromyalgia Treatment group was found to be statistically significantly higher than that of the Fibromyalgia group (p<0.001).

The relationships between EQIP, FKRE, and FKGL scores were examined. No statistically significant relationship was observed between EQIP and FKRE (p=0.156). A statistically significant negative relationship was observed between EQIP and FKGL values (r= -0.424; p<0.001). As FKRE increases, the FKGL value also significantly decreases, and the degree of relationship was determined to be moderate (r= -0.620; p<0.001).

DISCUSSION

In this study, the readability, comprehensibility, and quality of ChatGPT 3.5's answers to popular questions about fibromyalgia and fibromyalgia treatment were evaluated.

In their study Black et al. emphasized the gaps in knowledge and common misunderstandings about fibromyalgia among diagnosed individuals, highlighting the need for patient-centered educational interventions tailored to their experiences and understanding [10].

Also Daraz et al. identified the information needs and preferences of individuals living with fibromyalgia, highlighting the importance of reliable online resources and healthcare providers in addressing these needs [9]. Despite the evident need for online information sources, in several studies it was reported that websites providing information about fibromyalgia syndrome (FM) have poor content, low site quality, difficult readability [13,14]. ChatGPT 3.5 has become one of the most popular NLP models in recent times, widely utilized by numerous individuals. It possesses the potential to address the online information needs effectively. With its advanced capabilities in natural language processing, ChatGPT 3.5 stands out as a versatile tool for various tasks, including answering queries and providing valuable insights across a wide range of topics. This study suggests that ChatGPT 3.5 can be considered a high-quality source of information on keywords related to fibromyalgia and fibromyalgia treatment, despite its challenging readability.

Google Trends is a very effective method of determining how popular certain topics are and in which regions they receive more attention. These data show how the need for information about fibromyalgia and its treatment may vary geographically.

Analyzes of ChatGPT 3.5 answers present some important findings. In particular, FKRE and FKGL scores offer a critical evaluation in terms of readability and understandability of texts. According to the results of the study, it was determined that the texts were generally difficult to read, but the quality of information was high. This highlights ChatGPT's ability and quality of information to provide detailed information on complex topics like fibromyalgia in this case. In the literature, the usefulness of ChatGPT as a valuable source of information has been investigated in various medical topics [15-17]. This study is the first research on the usability of ChatGPT 3.5 in the context of fibromyalgia.

Research on the quality of responses provided by ChatGPT in chronic pain conditions is limited, with a primary focus on chronic low back pain [18,19]. Coraci et al. reported in their study that ChatGPT showed an acceptable and significant correlation with the Oswestry Disability Index and the Quebec Back Pain Disability Scale [18]. Also in this study, ChatGPT was found as a

source of accurate and understandable information. According to the studies of Coraci et al., it can be further concluded that, in addition to providing accurate information about a health problem, ChatGPT can also offer insights into the quality of life and disability levels associated with it.

Classification according to EQIP survey categories is also very important. This classification helps us understand which topics search terms focus on and what type of information users need most. For example, terms in the "Fibromyalgia" category generally focus on basic topics such as what the disease is, its symptoms, diagnosis, and genetic predisposition, while terms in the "Fibromyalgia Treatment" category point to more specific information such as treatment options, medications, diet, and natural treatments. However, in complex areas such as medications, while the readability of ChatGPT 3.5 is higher than other categories, its quality has been found to be lower. It is particularly noteworthy that the EQIP score of the terms in the Drug, medication or product group is lower than the other groups, and the EQIP score of the Discharge or aftercare group is high. This shows that the quality and understandability of health information may vary across specific topics and suggests that further improvements are needed in these areas. In studies which the quality of the answers given by ChatGPT was evaluated, such as this study, EQIP mean scores generally ranged between 40-50 [20-22]. Erden et al. [22], in their study investigating the information quality and readability of ChatGPT on osteoporosis, reported that readability differed significantly between EQIP categories that the readability of the Miscellaneous category is significantly higher than the Condition and illness. In this study, the readability index did not differ, but a significant difference was detected in the EQIP score showing the quality of the information, according to categories.

Additionally, when fibromyalgia and fibromyalgia treatment were compared, information about fibromyalgia was found to be easier to read, while information about fibromyalgia treatment was found to be more complex. No difference in quality was detected between the two subjects. This means that general information about fibromyalgia tends to be written in a more accessible and easier-to-understand manner, making it more readable for a broader audience. On the other hand, texts specifically focused on the treatment of fibromyalgia tend to be more complex, potentially requiring a higher level of understanding or expertise to fully comprehend.

One of the key advantages of ChatGPT directing readers to healthcare professionals in medical contexts is the assurance of accurate diagnosis and treatment. While AI can provide general information and support, it cannot replace the personalized care and expertise of a trained medical professional. This approach fosters a safer online environment by emphasizing the importance of professional oversight and reducing the risk of misinformation. Overall, the practice enhances patient safety and promotes better health outcomes.

CONCLUSIONS

Overall, this study provides an important look at how ChatGPT can be used in healthcare by examining in detail the success and limitations of ChatGPT in delivering health information.

In conclusion, while general fibromyalgia information provided by ChatGPT is generally more readable, the complexity of texts regarding fibromyalgia treatment is significantly higher, indicating a need for improved clarity and accessibility in treatment-related content to better serve patient understanding and care.

CRediT authorship contribution statement

Tugce Pasin: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Bilinc Dogruoz Karatekin: Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Ozge Pasin: Writing – review & editing, Formal analysis, Data curation, Conceptualization.

Table-I. Popular keywords according to Google Trends for Fibromyalgia and Fibromyalgia Treatment

Topic		Keyword	Topic Category for EQIP
Fibromyalgia	1	Fibromyalgia meaning	Condition or Ilness
	2	Fibromyalgia symptoms in women	Condition or Ilness
	3	Do I have Fibromyalgia	Condition or Ilness
	4	Is Fibromyalgia a disability	Miscellaneous
	5	What are symptoms of fibromyalgia	Condition or Ilness
	6	How is fibromyalgia diagnosed	Test, operation, investigation or procedure
	7	How to diagnose fibromyalgia	Test, operation, investigation or procedure
	8	Fibromyalgia symptoms in females	Condition or Ilness
	9	Fibromyalgia hereditary	Condition or Ilness
	10	Fibromyalgia symptoms list	Condition or Ilness
	11	Fibromyalgia signs and symptoms	Condition or Ilness
	12	Pregabalin	Drug, medication or product
	13	Duloxetine	Drug, medication or product
	14	Myalgia	Miscellaneous
	15	Naltrexone	Drug, medication or product
	16	Body ache	Miscellaneous
Fibromyalgia Treatment	1	What is fibromyalgia	Condition or Ilness
	2	Fibromyalgia syndrome	Condition or Ilness
	3	Fibromyalgia medication	Drug, medication or product
	4	Fibromyalgia treatment medication	Drug, medication or product
	5	Fibromyalgia cause	Condition or Ilness
	6	Best treatment for fibromyalgia	Discharge or aftercare
	7	Fibromyalgia diet	Discharge or aftercare
	8	Fibromyalgia test	Test, operation, investigation or procedure
	9	Fibromyalgia points	Test, operation, investigation or procedure
	10	Chronic pain	Miscellaneous
	11	New fibromyalgia treatment	Drug, medication or product
	12	What is fibromyalgia symptoms and treatment	Miscellaneous
	13	Natural treatment for fibromyalgia	Discharge or aftercare
	14	Fibromyalgia treatment guidelines	Drug, medication or product
	15	Fibromyalgia guidelines	Drug, medication or product
	16	Gabapentin	Drug, medication or product
	17	Fibromyalgia home treatment	Discharge or aftercare
	18	Pregabalin	Drug, medication or product
	19	Duloxetine	Drug, medication or product
	20	Alternative medicine	Discharge or aftercare
	21	Milnacipran	Drug, medication or product
	22	Amitriptyline	Drug, medication or product

Table-II. Descriptive Statistics for EQIP, FKRE, and FKGL scores

	Mean	Median	Std. Deviation	Minimum	Maximum
The Flesch-Kincaid Reading Ease Score	-2,9862	-3,8510	13,53061	-28,88	35,42
The Flesch-Kincaid Grade Level Score	19,1873	19,8960	2,36350	14,44	23,06
Ensuring Quality Information for Patients Score	41,6092	42,1500	9,18262	20,00	57,14

Table-III. Comparison of EQIP, FKRE, and FKGL scores among EQIP Categories

	EQIP Category	Mean	Median	Std. Deviation	Minimum	Maximum	p
The Flesch-Kincaid Reading Ease Score	Discharge or aftercare	-6,1682	-6,1320	9,22887	-16,28	4,11	0,300
	Condition or ilness	1,5462	3,3180	12,89097	-20,50	20,30	
	Test, operation, investigation or procedure	2,6618	4,8860	9,54771	-9,91	10,79	
	Drug, medication or product	-6,9621	-13,8800	16,90737	-28,88	35,42	
	Miscellaneous	-5,6586	-9,4440	11,58833	-16,38	10,53	
The Flesch-Kincaid Grade Level Score	Discharge or aftercare	19,1564	18,7850	1,65514	17,62	21,31	0,108
	Condition or ilness	17,6760	17,1425	2,56028	14,44	21,56	
	Test, operation, investigation or procedure	19,5905	19,4410	2,12962	17,42	22,07	
	Drug, medication or product	20,4817	20,6520	1,87781	16,95	23,06	
	Miscellaneous	19,4162	19,8470	2,38731	16,42	22,22	
Ensuring Quality Information for Patients Score	Discharge or aftercare	51,7820	49,9950	4,18886	48,21	57,14	0,004
	Condition or ilness	44,4967	45,2500	7,43326	34,30	56,25	
	Test, operation, investigation or procedure	46,8750	50,0000	8,98494	33,75	53,75	
	Drug, medication or product	34,9475	36,2500	7,53017	20,00	47,50	
	Miscellaneous	36,2820	35,9350	6,57902	26,79	45,30	

Table-IV. Comparison of EQIP, FKRE, and FKGL scores among Fibromyalgia and Fibromyalgia Keywords

	Keywords	Mean	Median	Std. Deviation	Minimum	Maximum	p
The Flesch-Kincaid Reading Ease Score	Fibromyalgia	4,7217	6,6145	10,14227	-17,06	20,30	0,002
	Fibromyalgia Treatment	-8,5920	-11,8575	13,07295	-28,88	35,42	
The Flesch-Kincaid Grade Level Score	Fibromyalgia	17,2584	17,1200	2,05748	14,44	22,22	<0,001
	Fibromyalgia Treatment	20,5902	20,6820	1,38630	17,62	23,06	
Ensuring Quality Information for Patients Score	Fibromyalgia	44,1597	46,4000	7,95098	33,75	56,25	0,147
	Fibromyalgia Treatment	39,7543	39,0600	9,73902	20,00	57,14	

REFERENCES

1. Seol HY, Rolfes MC, Chung W, et al. Expert artificial intelligence-based natural language processing characterises childhood asthma. *BMJ open respiratory research*. Feb 2020;**7**(1).

2. Sivarajkumar S, Gao F, Denny P, et al. Mining Clinical Notes for Physical Rehabilitation Exercise Information: Natural Language Processing Algorithm Development and Validation Study. *JMIR medical informatics*. Oct 2024;**12**:e52289.

3. Nazir A, Wang Z. A Comprehensive Survey of ChatGPT: Advancements, Applications, Prospects, and Challenges. Meta-radiology. Sep 2023;1(2).

4. Ray PP. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*. 2023;**3**:121-154.

5. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Frontiers in artificial intelligence*. 2023;**6**:1169595.

6. Qiu J, Li L, Sun J, et al. Large AI Models in Health Informatics: Applications, Challenges, and the Future. IEEE journal of biomedical and health informatics. Dec 2023;27(12):6074-6087.

7. Temsah MH, Aljamaan F, Malki KH, et al. ChatGPT and the Future of Digital Health: A Study on Healthcare Workers' Perceptions and Expectations. *Healthcare* (Basel, Switzerland). Jun 21 2023;**11**(13).

8. Javaid M, Haleem A, Singh RP. ChatGPT for healthcare services: An emerging stage for an innovative perspective. BenchCouncil Transactions on Benchmarks, *Standards and Evaluations*. 2023/02/01/ 2023;**3**(1):100105.

9. Daraz L, Macdermid J, Wilkins S, Gibson J, Shaw L. Information preferences of people living with fibromyalgia – a survey of their information needs and preferences. Rheumatology Reports. 01/13 2011;3.

10. Black LL, Black WR, Chadwick A, Christofferson JL, Katz H, Kragenbrink M. Investigation of patients' understanding of fibromyalgia: Results from an online qualitative survey. Patient Education and Counseling. 2024/05/01/ 2024;122:108156.

11. Kincaid P, Fishburne RP, Rogers RL, Chissom BS. Derivation of New Readability Formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy Enlisted Personnel. 1975

12. Moult 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 expectations : *an international journal of public participation in health care and health policy*. Jun 2004;**7**(2):165-75.

13. Otu M, Karagözoğlu Ş. Investigating the Websites in Turkey that Providing Information on Fibromyalgia Syndrome by Readability, Content and Quality. *Turkish Journal of Osteopo*rosis. 2022;**28**:19-25.

14. Alioshkin Cheneguin A, Salvat Salvat I, Romay Barrero H, Torres Lacomba M. How good is online information on fibromyalgia? An analysis of quality and readability of websites on fibromyalgia in Spanish. *BMJ open*. Jul 5 2020;**10**(7):e037065.

15. Cinar C. Analyzing the Performance of ChatGPT About Osteoporosis. *Cureus*. Sep 2023;**15**(9):e45890.

16. Choi J, Kim JW, Lee YS, et al. Availability of ChatGPT to provide medical information for patients with kidney cancer. *Scientific Reports*. 2024/01/17 2024;**14**(1):1542.

17. Ozgor BY, Simavi MA. Accuracy and reproducibility of ChatGPT's free version answers about endometriosis. International journal of gynaecology and obstetrics: the official organ of the *International Federation of Gynaecology and Obstetrics*. May 2024;**165**(2):691-695.

18. Coraci D, Maccarone MC, Regazzo G, Accordi G, Papathanasiou JV, Masiero S. ChatGPT in the development of medical questionnaires. The example of the low back pain. *European journal of translational myology*. Dec 15 2023;**33**(4).

19. Narula P, Deoghare R. Evaluation of ChatGPT responses for back pain? European journal of orthopaedic surgery & traumatology : *orthopedie traumatologie*. Apr 2024;**34**(3):1737-1739.

20. Yurdakurban E, Topsakal KG, Duran GS. A comparative analysis of AI-based chatbots: Assessing data quality in orthognathic surgery related patient information. *Journal of stomatology, oral and maxillofacial surgery*. Dec 28 2023;**125**(5):101757.

21. Temel MH, Erden Y, Bağcıer F. Information Quality and Readability: ChatGPT's Responses to the Most Common Questions About Spinal Cord Injury. *World neurosurgery*. Jan 2024;**181**:e1138-e1144.

22. Erden Y, Temel MH, Bağcıer F. Artificial intelligence insights into osteoporosis: assessing ChatGPT's information quality and readability. *Archives of osteoporosis*. Mar 19 2024;**19**(1):17.