

How reliable is the artificial intelligence product large language model ChatGPT in orthodontics?

Kevser Kurt Demirsoy^a; Suleyman Kutalmış Buyuk^b; Tayyip Bicer^c

ABSTRACT

Objectives: To evaluate the reliability of information produced by the artificial intelligence-based program ChatGPT in terms of accuracy and relevance, as assessed by orthodontists, dental students, and individuals seeking orthodontic treatment.

Materials and Methods: Frequently asked and curious questions in four basic areas related to orthodontics were prepared and asked in ChatGPT (Version 4.0), and answers were evaluated by three different groups (senior dental students, individuals seeking orthodontic treatment, orthodontists). Questions asked in these basic areas of orthodontics were about: clear aligners (CA), lingual orthodontics (LO), esthetic braces (EB), and temporomandibular disorders (TMD). The answers were evaluated by the Global Quality Scale (GQS) and Quality Criteria for Consumer Health Information (DISCERN) scale.

Results: The total mean DISCERN score for answers on CA for students was 51.7 ± 9.38 , for patients was 57.2 ± 10.73 and, for orthodontists was 47.4 ± 4.78 ($P = .001$). Comparison of GQS scores for LO among groups: students (3.53 ± 0.78), patients (4.40 ± 0.72), and orthodontists (3.63 ± 0.72) ($P < .001$). Intergroup comparison of ChatGPT evaluations about TMD was examined in terms of the DISCERN scale, with the highest value given in the patients group (57.83 ± 11.47) and lowest value in the orthodontist group (45.90 ± 11.84). When information quality evaluation about EB was examined, it GQS scores were >3 in all three groups (students: 3.50 ± 0.78 ; patients: 4.17 ± 0.87 ; orthodontists: 3.50 ± 0.82).

Conclusions: ChatGPT has significant potential in terms of usability for patient information and education in the field of orthodontics if it is developed and necessary updates are made. (*Angle Orthod.* 2024;94:602–607.)

KEY WORDS: ChatGPT; Clear aligners; Lingual orthodontics; Artificial intelligence; Patient information

INTRODUCTION

Artificial intelligence (AI) has emerged as a transformative force in various fields of healthcare, offering innovative solutions to complex challenges.¹ In orthodontics, AI

presents exciting opportunities for improving patient education and information dissemination.² ChatGPT, a cutting-edge language model powered by AI, has garnered significant attention for its ability to generate coherent and contextually relevant responses to user inquiries.³ ChatGPT is, capable of producing creative and contextually relevant responses, but it does not search the internet in real time like Google's search engine. However, the reliability and utility of ChatGPT in the orthodontic context remain a subject of investigation.⁴

AI-based robots are also commonly used for dental education among students in the field, serving as easily accessible resources. Therefore, the accuracy and reliability of information in this domain are also crucial from an educational standpoint.⁵ Patients seeking orthodontic treatment or experiencing various joint disorders may seek information on these subjects through various online sources.⁶ In this process, it is essential for orthodontists or dental specialists to provide clear and

^a Associate Professor, Department of Orthodontics, Faculty of Dentistry, Nevsehir Hacı Bektas Veli University, Nevsehir, Turkey.

^b Associate Professor, Department of Orthodontics, Faculty of Dentistry Ordu University, Ordu, Turkey.

^c Research Assistant, Department of Orthodontics, Faculty of Dentistry Ordu University, Ordu, Turkey.

Corresponding author: Dr Kevser Kurt Demirsoy, Department of Orthodontics, Faculty of Dentistry Nevsehir Hacı Bektas Veli University, Nevsehir, Turkey (e-mail: k_idemirsoy@hotmail.com)

Accepted: June 2024. Submitted: March 2024.

Published Online: August 14, 2024

© 2024 by The EH Angle Education and Research Foundation, Inc.

concise responses to patient inquiries and to prepare them for the treatment journey. However, it is not always feasible for patients to consult orthodontists or dentists to address their concerns regarding the process. AI-based chatbots hold significant potential to offer relevant information to patients.⁷ There are no studies evaluating the accuracy and reliability of ChatGPT in ensuring patient information and education about various orthodontic treatments across different societal units.

This study aimed to assess the credibility of ChatGPT responses to orthodontic queries, focusing on its potential for enhancing patient education and information dissemination. By evaluating perceptions of quality among orthodontists, dental students, and orthodontic patients, this research sought to shed light on the strengths and limitations of ChatGPT in the orthodontic domain. Through a structured analysis of responses across different orthodontic topics including clear aligners (CA), lingual orthodontics (LO), esthetic braces (EB), and temporomandibular disorders (TMD), this study aimed to inform regarding strategies for optimizing the utility of ChatGPT in clinical practice.

MATERIALS AND METHODS

This study received ethical approval from the Nevsehir Haci Bektas Veli University Non-Interventional Research Ethics Board (Report No: 2023.09.08).

Study Design and Survey Questions

In this study, commonly asked and curious questions related to four fundamental areas of orthodontics were prepared and directed to the artificial intelligence-based chatbot ChatGPT, and the responses obtained were evaluated by three different groups. Each group provided feedback on the same set of responses, allowing for comprehensive evaluation of the perceived accuracy and relevance of ChatGPT-generated information within different user demographics. These four fundamental areas of orthodontics included:

1. Can you provide information about clear aligners?
2. Can you provide information about lingual orthodontics?
3. Can you provide information about esthetic braces?
4. Can you provide information about temporomandibular disorders?

Number and Attributes of Participants

ChatGPT responses were evaluated by three different groups:

1. Orthodontists: This group consisted of experienced orthodontists specializing in their field. They are professionals who specialize in orthodontics and typically have clinical experience.

2. Fifth-year dental students: This group was comprised of students in the final years of their dental education. These students have theoretical knowledge as well as clinical experience.
3. Individuals aged 18 years and older who expressed a desire for orthodontic treatment: Participants in this group were individuals aged 18 and above who required orthodontic treatment. These individuals typically seek relevant sources to obtain information about their own orthodontic conditions or to learn about treatment options.

The sample size was calculated using the G*Power program (version 3.1.9.2; Axel Buchner, University of Düsseldorf, Düsseldorf, Germany) to detect a medium effect size with 90% power, requiring a total sample size of 98 (effect size: 0.60). To increase data reliability, it was planned for a total of 105 individuals to participate, with 35 individuals in each group.

Assessment Scales (DISCERN and GQS)

To assess the reliability of the responses provided by Chat GPT, the DISCERN tool was used, and the overall quality of information was evaluated using the Global Quality Scale (GQS). The Quality Criteria for Consumer Health Information (DISCERN) was developed to enable patients and healthcare providers to assess the quality of health information.^{8,9} The DISCERN scale consists of 16 questions, each rated on a scale of 1 (low quality) to 5 (high quality). The first eight questions assessed reliability, the next seven questions evaluated treatment options, and the final question assessed the overall quality of the information based on the responses to the first 15 questions.⁹ The GQS, consisting of five questions, was used as a second tool to assess the quality and usefulness of the provided information for patients. Scores were calculated by totaling the points for each section. Videos with a total GQS score ≤ 3 were classified as being between low and moderate quality, while videos with a score > 3 were classified as being between good and excellent quality.^{10,11} These tools allowed for a structured assessment of factors such as clarity, accuracy, and relevance of the information provided by ChatGPT.

Statistical Analysis

Statistical analyses were conducted using SPSS software (SPSS Inc., Statistical Package for Social Sciences, version 20.0, Chicago, IL, USA). The normality of data distribution was assessed using the Shapiro-Wilk test, while demographic characteristics were compared among groups using the Kruskal-Wallis test. The Mann-Whitney *U*-test was employed to compare GQS and DISCERN parameters among groups. Spearman's

Table 1. ChatGPT Results for Clear Aligner Systems. Comparison of DISCERN and GQS Scores Among Groups

Parameters	Senior Dentistry Students (n = 30)		Individuals Seeking Orthodontic Treatment (n = 30)		Orthodontists (n = 30)		P*
	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	
DISCERN							
Section 1	27.33 (5.08)	19–40	30.53 (5.58)	20–40	25.53 (4.55)	16–40	.002
Section 2	24.37 (4.96)	16–35	26.67 (5.79)	15–35	21.87 (4.55)	14–35	.005
Total Mean	51.7 (9.38)	35–75	57.2 (10.73)	36–75	47.4 (4.78)	31–75	.001
Section 3	3.57 (0.90)	1–5	3.97 (0.89)	2–5	3.47 (0.73)	2–5	.078
GQS	3.43 (0.77)	2–5	4.27 (0.69)	3–5	3.67 (0.76)	2–5	<.001

^a GQS indicates Global Quality Scale; Max, Maximum; Min, Minimum; SD, standard deviation.

* Results of Kruskal-Wallis test. Bold values are statistically significant.

correlation coefficients were calculated to evaluate potential correlations between DISCERN and GQS parameters. Statistical significance was set at $P < .05$.

RESULTS

Analysis of ChatGPT responses revealed varying perceptions of reliability among participant groups and across different orthodontic topics. Overall, responses were rated as medium to good quality, indicating a reasonable level of accuracy and clarity in the information provided by ChatGPT. However, significant differences were observed in evaluations among orthodontists, dental students, and orthodontic patients.

Comparison of DISCERN and GQS scores of ChatGPT results for Clear Aligner Systems among groups is presented in Table 1. Overall, significant differences were observed among the participant groups in their assessment of ChatGPT responses about CA ($P^* < .05$). In terms of the DISCERN scale, individuals seeking orthodontic treatment rated the responses highest, with a total mean score of 57.2 (SD: 10.73), followed by senior dental students with a total mean score of 51.7 (SD: 9.38), and orthodontists with a total mean score of 47.4 (SD: 4.78). The GQS scores also showed significant differences among participant groups ($P < .001$). Individuals seeking orthodontic treatment rated the responses highest on the GQS (mean: 4.27, SD: 0.69), followed by senior dental students (mean: 3.43, SD: 0.77) and orthodontists (mean: 3.67, SD: 0.76).

Table 2 presents a comparison of DISCERN and GQS scores for ChatGPT responses related to lingual orthodontics among the groups. Significant differences were observed among the participant groups in their assessments of ChatGPT responses related to LO ($P^* < .05$). Individuals seeking orthodontic treatment rated the responses highest on both the DISCERN scale and GQS, with a total mean DISCERN score of 58.50 and a GQS score of 4.40.

Table 3 compares the DISCERN and GQS scores for ChatGPT responses related to TMD among the groups. Significant differences were observed among the participant groups in their assessments of ChatGPT responses related to TMD ($P^* < .05$). Individuals seeking orthodontic treatment rated the responses highest on both the DISCERN scale and GQS, with a total mean DISCERN score of 57.83 and a GQS score of 4.27. This was followed by senior dental students, who had a total mean DISCERN score of 49.63 and a GQS score of 3.53, and orthodontists, who had a total mean DISCERN score of 45.90 and a GQS score of 3.43.

Table 4 displays the comparison of DISCERN and GQS scores for ChatGPT responses related to esthetic orthodontic braces among the groups. Significant differences were observed among the participant groups in their assessments of ChatGPT responses related to EB ($P^* < .05$). Individuals seeking orthodontic treatment rated the responses highest on both the DISCERN scale and GQS, with a total mean DISCERN score of 58.33 and a GQS score of 4.17. This was followed by senior

Table 2. ChatGPT Results for Lingual Orthodontics. Comparison of DISCERN and GQS Scores Among Groups

Parameters	Senior Dentistry Student (n = 30)		Individuals Seeking Orthodontic Treatment (n = 30)		Orthodontists (n = 30)		P*
	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	
DISCERN							
Section 1	26.43 (6.40)	15–40	31.13 (5.80)	22–40	24.93 (5.32)	13–40	.001
Section 2	25.10 (5.16)	14–35	27.37 (4.92)	19–35	22.87 (5.19)	11–35	.005
Total Mean	51.53 (11.16)	29–75	58.50 (10.55)	43–75	47.8 (9.09)	24–75	.001
Section 3	3.67 (0.96)	1–5	4.03 (0.81)	3–5	3.63 (0.76)	2–5	.159
GQS	3.53 (0.78)	2–5	4.40 (0.72)	3–5	3.63 (0.72)	2–5	<.001

^a GQS indicates Global Quality Scale; Max, Maximum; Min, Minimum; SD, standard deviation.

* Results of Kruskal-Wallis test. Bold values are statistically significant.

Table 3. ChatGPT Results for Temporomandibular Disorders. Comparison of DISCERN and GQS Scores Among Groups

Parameters	Senior Dentistry Student (n = 30)		Individuals Seeking Orthodontic Treatment (n = 30)		Orthodontists (n = 30)		<i>P</i> *
	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	
DISCERN							
Section 1	26.13 (6.34)	9–40	30.60 (6.29)	21–40	24.40 (6.23)	10–40	.002
Section 2	23.50 (4.92)	13–35	27.23 (5.34)	19–35	21.50 (6.28)	10–35	.002
Total Mean	49.63 (10.81)	29–75	57.83 (11.47)	40–75	45.90 (11.84)	20–75	.001
Section 3	3.57 (0.81)	2–5	3.97 (0.85)	3–5	3.47 (1.04)	1–5	.155
GQS	3.53 (0.86)	2–5	4.27 (0.87)	2–5	3.43 (0.90)	1–5	.001

^a GQS indicates Global Quality Scale; Max, Maximum; Min, Minimum; SD, standard deviation.

* Results of Kruskal-Wallis test. Bold values are statistically significant.

dental students, who had a total mean DISCERN score of 52.37 and a GQS score of 3.50, and orthodontists, who had a total mean DISCERN score of 47.70 and a GQS score of 3.50.

DISCUSSION

AI-driven chatbots are advanced software programs designed to engage in human-like conversations, employing natural language interfaces to deliver a wide range of services through text-based interactions.^{12–14} These intelligent bots serve diverse functions, including disseminating information on specific subjects, addressing inquiries, offering customer support, and simulating therapeutic dialogue with users.^{15–17} The development of such bots heavily relies on AI methodology, leveraging techniques like natural language processing, machine learning, and deep learning to enrich their linguistic comprehension and interactivity.¹⁸ Notably, recent research has showcased the remarkable versatility of ChatGPT, a prominent chatbot model capable of autonomously generating scientific articles.¹⁹

In recent years, an increasing number of adult patients have been seeking orthodontic treatment and exploring alternatives to traditional fixed appliances for esthetic and comfort reasons.²⁰ The use of clear aligners as orthodontic appliances emerged in 1946 when Kesling designed a series of thermoplastic tooth positioners to gradually move misaligned teeth to better positions.²¹ In 1997, Align

Technology (Santa Clara, Calif) adapted and merged modern technologies to introduce clear aligner therapy (CAT), making Kesling’s concept a feasible orthodontic treatment option.²² During the evaluation of ChatGPT responses regarding Clear Aligner treatment, all participant groups provided positive ratings. However, individuals seeking orthodontic treatment consistently rated the responses higher compared to senior dental students and orthodontists, suggesting a higher level of satisfaction and perceived quality of information provided by ChatGPT.

Lingual orthodontics and esthetic braces have also emerged as esthetic alternatives to conventional orthodontic treatment. Many adult patients seek information on these topics online. In terms of specific sections of the DISCERN scale, individuals seeking orthodontic treatment consistently rated the reliability of information (Section 1) and the quality of information on treatment choices about lingual orthodontics (Section 2) highest compared to the other participant groups. Orthodontists consistently rated the responses lowest across both sections. Overall, the findings suggested that individuals seeking orthodontic treatment had the most positive perceptions of ChatGPT responses related to Lingual Orthodontics, followed by senior dental students and orthodontists. In a manner similar to that observed for lingual orthodontics, individuals seeking orthodontic treatment consistently rated the reliability of information provided by ChatGPT regarding esthetic braces highest compared to the other participant groups.

Table 4. ChatGPT Results for Esthetic Orthodontic Braces. Comparison of DISCERN and GQS Scores Among Groups

Parameters	Senior Dentistry Students (n = 30)		Individuals Seeking Orthodontic Treatment (n = 30)		Orthodontists (n = 30)		<i>P</i>
	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	Mean (SD)	Min.–Max.	
DISCERN							
Section 1	27.47 (5.58)	17–38	30.97 (6.68)	18–40	24.70 (5.40)	13–40	<.001**
Section 2	24.90 (5.82)	11–34	27.37 (5.49)	18–35	23.00 (5.21)	13–35	.015*
Total Mean	52.37 (11.01)	30–72	58.33 (12.04)	37–75	47.70 (9.82)	26–75	.001**
Section 3	3.67 (0.96)	1–5	4.07 (0.91)	2–5	3.57 (0.82)	2–5	.076*
GQS	3.50 (0.78)	2–5	4.17 (0.87)	2–5	3.50 (0.82)	2–5	.003*

^a GQS indicates Global Quality Scale; Max, Maximum; Min, Minimum; SD, standard deviation.

* Results of Kruskal-Wallis test. Bold values are statistically significant.

** Results of one-way analysis of variance.

TMD is a clinical condition characterized by pain, sounds (crepitus or clicking), and irregular movements in the temporomandibular joint, posing one of the most challenging conditions to treat among causes of maxillofacial pain.²³ TMD is common, with a prevalence ranging from 5% to 16%. It is more prevalent in women, with an incidence that increases with age.²⁴ TMD has been described using a wide range of terminology. However, in 1982, the American Dental Association recommended the clear diagnostic and therapeutic distinction of various conditions affecting TMD and masticatory muscles and preferred the definition of TMD. These disorders can cause serious problems as they affect functions such as eating, speaking, breathing, and swallowing, highlighting the importance of their treatment.²⁵ When evaluating the information provided by ChatGPT on TMD across the three different groups using specific sections of the DISCERN scale, individuals seeking orthodontic treatment consistently rated the reliability of information (Section 1) and the quality of information on treatment choices (Section 2) higher compared to the other participant groups. Orthodontists consistently rated the responses lowest across both sections. Overall, the findings suggest that individuals seeking orthodontic treatment have the most positive perceptions of ChatGPT responses related to TMD, followed by senior dental students and orthodontists.

AI-based robots can also be frequently used for dental students in education and are among easily accessible resources. Therefore, the accuracy and reliability of the information in this field is also important for educational purposes. Patients seeking orthodontic treatment or experiencing various joint disorders may want to obtain information on these topics through various internet sources. During this process, it is crucial for orthodontists or dental specialists to provide clear and concise answers to patient questions and prepare them for the treatment process. However, it is not always possible for patients to consult orthodontists or dentists to address their concerns about the process. Artificial intelligence-based chatbots have tremendous potential to provide relevant information to patients. Presently, there remains a gap in research concerning the efficacy and precision of ChatGPT in delivering orthodontic treatment information and education across diverse social contexts.

The findings of this study offered valuable insights into the reliability and utility of ChatGPT in the orthodontic context. While ChatGPT demonstrates potential as an educational tool for patients and practitioners alike,⁶ several factors must be considered to optimize its effectiveness in clinical practice.^{3,5}

One key consideration is the need for ongoing refinement and customization of ChatGPT responses to better meet the diverse needs of orthodontic stakeholders. Tailoring responses to specific patient demographics,

treatment modalities, and clinical scenarios may enhance the relevance and accuracy of the information provided.²⁶ Additionally, integrating feedback mechanisms to continuously improve the performance of ChatGPT based on real-world usage data could further enhance its utility in orthodontic practice.

In addition, efforts to enhance transparency and accountability in AI-generated information are essential to build trust among users and ensure responsible deployment of AI technology in healthcare.²⁷ Providing users with clear indications of the limitations and uncertainties associated with AI-generated responses can help manage expectations and promote informed decision-making.^{28–30} ChatGPT is a language model designed to generate text based on the input it receives, but it doesn't guarantee the accuracy or veracity of the text it produces. Therefore, it is not advisable to rely solely on the information generated by ChatGPT without verification. Researchers and professionals should carefully evaluate content generated by ChatGPT and verify it when necessary.

There were some limitations of this study.

Sample Composition

The study sample primarily consisted of orthodontists, senior dental students, and individuals seeking orthodontic treatment. While these groups represented key stakeholders in orthodontic care, the study may benefit from including a more diverse range of participants, such as general dentists, other healthcare professionals, and individuals with different demographic backgrounds.

Generalizability

The findings of this study may have limited generalizability beyond the specific context and population examined. Variations in patient preferences, cultural factors, and healthcare systems could influence the perceptions of the reliability and utility of ChatGPT in different settings.

Single Assessment Method

The reliability and utility of ChatGPT responses were evaluated primarily using the DISCERN scale and GQS. While these tools provided valuable insights into the quality of information provided, additional assessment methods, such as qualitative interviews or usability testing, could offer complementary perspectives on the performance of ChatGPT.

CONCLUSIONS

- This study underscores the potential of ChatGPT as a valuable resource for patient education and information dissemination in orthodontics.

- While ChatGPT demonstrates reasonable reliability in providing orthodontic information, ongoing refinement and customization are essential to optimize its effectiveness in clinical practice.
- With continued improvements in the future, orthodontists may be able to use such AI systems to optimize treatment outcomes.

ACKNOWLEDGMENTS

The data underlying this study are available upon request from the corresponding author.

DISCLOSURES

The authors declare that they have no competing interests.

REFERENCES

1. Yu KH, Beam AL, Kohane IS. Artificial intelligence in health-care. *Nat Biomed Eng*. 2018;2(10):719–731.
2. Kielczykowski M, Kamiński K, Perkowski K, Zadurska M, Czochrowska E. Application of Artificial Intelligence (AI) in a Cephalometric Analysis: A Narrative Review. *Diagnostics (Basel)*. 2023;13(16):2640.
3. Nazir A, Wang Z. A Comprehensive Survey of ChatGPT: Advancements, Applications, Prospects, and Challenges. *Meta Radiol*. 2023;1(2):100022.
4. Hwang JJ, Jung YH, Cho BH, Heo MS. An overview of deep learning in the field of dentistry. *Imaging Sci Dent*. 2019;49(1):1–7.
5. Dipalma G, Inchingolo AD, Inchingolo AM, et al. Artificial Intelligence and Its Clinical Applications in Orthodontics: A Systematic Review. *Diagnostics (Basel)*. 2023;13(24):3677.
6. Bichu YM, Hansa I, Bichu AY, Premjani P, Flores-Mir C, Vaid NR. Applications of artificial intelligence and machine learning in orthodontics: a scoping review. *Prog Orthod*. 2021;22(1):18.
7. Alzaid N, Ghulam O, Albani M, et al. Revolutionizing Dental Care: A Comprehensive Review of Artificial Intelligence Applications Among Various Dental Specialties. *Cureus*. 2023;15(10):e47033.
8. 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;53:105–111.
9. Meade MJ, Dreyer CW. Web-based information on orthodontic clear aligners: a qualitative and readability assessment. *Aust Dent J*. 2020;65:225–232.
10. Livas C, Delli K, Ren Y. Quality evaluation of the available Internet information regarding pain during orthodontic treatment. *Angle Orthod*. 2013;83:500–506.
11. 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(1-2):31–39.
12. Aggarwal A, Tam CC, Wu D, Li X, Qiao S. Artificial intelligence-based chatbots for promoting health behavioral changes: systematic review. *J Med Internet Res*. 2023;25:e40789.
13. Wailthare S, Gaikwad T, Khadse K, Dubey P. Artificial intelligence based chat-bot. *Artif Intell*. 2018;5.
14. Krishnan C, Gupta A, Gupta A, Singh G. Impact of artificial intelligence-based chatbots on customer engagement and business growth. In: *Deep Learning for Social Media Data Analytics*. Springer; 2022:195–210.
15. Verma S, Sharma R, Deb S, Maitra D. Artificial intelligence in marketing: systematic review and future research direction. *Int J Inf Manag Data Insights*. 2021;1:100002.
16. Vlačić B, Corbo L, e Silva SC, Dabić M. The evolving role of artificial intelligence in marketing: a review and research agenda. *J Bus Res*. 2021;128:187–203.
17. Arsenijevic U, Jovic M. Artificial intelligence marketing: chatbots. International Conference on Artificial Intelligence. Applied Innovations, IEEE; 2019:19–193.
18. Matic R, Kabiljo M, Zivkovic M, Cabarkapa M. Extensible chatbot architecture using metamodels of natural language understanding. *Electronics*. 2021;10:2300.
19. Benichou L; ChatGPT. The role of using ChatGPT AI in writing medical scientific articles. *J Stomatol Oral Maxillofac Surg*. 2023;124(5):101456.
20. Rosvall MD, Fields HW, Ziuchkovski J, Rosenstiel SF, Johnston WM. Attractiveness, acceptability, and value of orthodontic appliances. *Am J Orthod Dentofacial Orthop*. 2009;135:276.
21. Kesling HD. Coordinating the predetermined pattern and tooth positioner with conventional treatment. *Am J Orthod Oral Surg*. 1946;32:285–293.
22. Rossini G, Parrini S, Castroflorio T, Deregibus A, Debernardi CL. Efficacy of clear aligners in controlling orthodontic tooth movement: a systematic review. *Angle Orthod*. 2015;85(5):881–889.
23. Griffiths RH. Report of the President's conference on examination, diagnosis and management of temporomandibular disorders. *J Am Dent Assoc*. 1983;106(1):75–77.
24. Cahlin BJ, Dahlström L. No effect of glucosamine sulfate on osteoarthritis in the temporomandibular joints—a randomized, controlled, short-term study. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2011;112(6):760–766.
25. Nicolakis P, Erdogmus B, Kopf A, et al. Effectiveness of exercise therapy in patients with internal derangement of the temporomandibular joint. *J Oral Rehabil*. 2001;28(2):1158–1164.
26. Thurzo A, Urbanová W, Novák B, et al. Where Is the Artificial Intelligence Applied in Dentistry? Systematic Review and Literature Analysis. *Healthcare (Basel)*. 2022;10(7):1269.
27. Baker A, Perov Y, Middleton K, et al. A Comparison of Artificial Intelligence and Human Doctors for the Purpose of Triage and Diagnosis. *Front Artif Intell*. 2020;3:543405.
28. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*. 2019;25(1):44–56.
29. Liu J, Zhang C, Shan Z. Application of Artificial Intelligence in Orthodontics: Current State and Future Perspectives. *Healthcare (Basel)*. 2023;11(20):2760.
30. Mohaideen K, Negi A, Verma DK, Kumar N, Sennimalai K, Negi A. Applications of artificial intelligence and machine learning in orthognathic surgery: A scoping review. *J Stomatol Oral Maxillofac Surg*. 2022;123(6):e962–e972.