

Academia Journal of Medicine

Year 2026, Volume-9, Issue- 1 (January- June)



Assessment of ChatGPT-Generated Endodontic Patient Education Materials: A Readability and Accuracy Analysis

¹Dr. Perisetty Dinesh Kumar, ²Dr. Shrenik Nahata, ³Dr. Sumit Bhatt, ⁴Dr. Sai Sagar, ⁵Dr. Milind Rajan

¹Reader, Department of Conservative Dentistry & Endodontics, Narayana Dental college & Hospital, Chintha Reddy Palem, Nellore, Andhra Pradesh 524002.

²Senior Lecturer, Department of Oral Pathology and Microbiology, Rungta College of Dental Sciences and Research, Bhilai, Chhattisgarh.

³PhD Scholar, Assistant Professor, Department of Oral and Maxillofacial Surgery, Rajasthan Dental College & Hospital, Nirwan University, Jaipur, Rajasthan.

⁴Senior House Surgeon, Govt Hospital Neyattinkara, Trivandrum, Kerala.

⁵PhD Scholar, Department of Pediatric Dentistry, Sharad Pawar Dental College and Hospital, Datta Meghe Institute of Higher Education and Research, Wardha-442107.

ARTICLE INFO

Keywords:

ChatGPT,
Artificial Intelligence,
Endodontics,
Patient Education,
Readability

doi:10.48165/ajm.2026.9.01.43

ABSTRACT

Background: The increasing use of artificial intelligence-based conversational models such as ChatGPT has transformed the way patients access healthcare-related information. In endodontics, patients frequently seek online educational material regarding root canal treatment and related conditions. However, the readability and scientific accuracy of AI-generated dental information remain uncertain.

Aim: To evaluate the readability and scientific accuracy of ChatGPT-generated patient education materials related to common endodontic conditions and procedures.

Materials and Methods: A cross-sectional analytical study was conducted using 100 patient-oriented responses generated by ChatGPT. Ten standardized endodontic prompts related to root canal treatment, dental abscess, infected teeth, post-treatment care, and treatment complications were entered ten separate times under identical settings. The generated responses were collected without modification. Readability assessment was performed using the Flesch Reading Ease Score, Flesch-Kincaid Grade Level, SMOG Index, and Gunning Fog Index. Scientific accuracy and quality of information were independently evaluated by three experienced endodontists using a structured 5-point Likert scale assessing accuracy, completeness, clarity, misleading information, and clinical usefulness. Statistical analysis was performed using descriptive statistics and Cohen's kappa test for inter-examiner reliability.

Results: ChatGPT-generated educational materials demonstrated moderate to high readability, with most responses corresponding to a secondary school reading level. The majority of responses showed good scientific accuracy, clarity, and clinical usefulness. However, certain responses lacked comprehensive explanations regarding procedural complications and emergency management. Minor inconsistencies and oversimplified information were observed in selected outputs. Inter-examiner agreement among evaluators was high.

Conclusion: ChatGPT demonstrated promising potential in generating readable and reasonably accurate endodontic patient education materials. Nevertheless, professional verification by dental experts remains necessary before clinical use to ensure reliability and prevent dissemination of incomplete or misleading information.

Introduction

Artificial intelligence (AI) has rapidly emerged as a transformative technology in healthcare, significantly influencing the way medical and dental information is accessed, processed, and communicated.^{1,2} Among the recent advancements in AI, conversational language models such as ChatGPT have gained considerable popularity due to their ability to generate human-like responses to user queries. These tools are increasingly being utilized by the general population to obtain healthcare-related information, including explanations of diseases, treatment procedures, postoperative instructions, and preventive care measures.^{3,4} In dentistry, patients frequently rely on internet-based resources to seek information regarding oral diseases and dental treatments before consulting a clinician. Endodontic procedures, particularly root canal treatment, are commonly associated with fear, anxiety, and misconceptions among patients.^{5,6} Lack of adequate understanding regarding treatment procedures and outcomes may negatively influence patient compliance and decision-making. Therefore, the availability of accurate, understandable, and reliable patient education material plays an essential role in improving oral health literacy and reducing treatment-related anxiety.⁷

Traditional online dental information sources often vary in readability, scientific quality, and reliability. Recently, AI-based conversational systems have emerged as alternative tools capable of providing instant and simplified explanations tailored to patient queries. Despite their growing use, concerns remain regarding the accuracy, completeness, and potential dissemination of misleading information generated by these platforms.⁸ In healthcare communication, readability is an important factor, as educational materials written at excessively high reading levels may not be easily understood by the general population. Similarly, inaccurate or incomplete information may adversely affect patient understanding and clinical decision-making.⁹

Several studies in the medical field have evaluated the performance of AI-generated educational content; however, limited literature exists regarding the quality of AI-generated patient information in endodontics. Assessing the readability and scientific accuracy of such content is therefore necessary before its integration into routine patient education.

Hence, the present study was conducted to evaluate the readability and scientific accuracy of ChatGPT-generated patient education materials related to common endodontic conditions and procedures.

Material and Methods

Study Design: A cross-sectional analytical study was conducted to evaluate the readability and scientific accuracy

of patient education materials generated using ChatGPT for common endodontic conditions and procedures.

Study Population: The study included 100 patient-oriented responses generated by ChatGPT. Ten standardized endodontic prompts were entered ten separate times under identical settings, resulting in a total of 100 responses for analysis.

Selection of Topics and Standardized Prompts: Frequently searched endodontic topics were identified based on common patient concerns and online search trends. Standardized patient-oriented prompts were prepared in simple English language and entered into ChatGPT under identical settings.

The prompts used in the study were as follows:

1. What is root canal treatment?
2. Why is root canal treatment needed?
3. Is root canal treatment painful?
4. What are the symptoms of an infected tooth?
5. What are the symptoms of a dental abscess?
6. How long does root canal treatment take?
7. What should I do after root canal treatment?
8. What foods should be avoided after root canal treatment?
9. What are the complications of root canal treatment?
10. What happens if I delay root canal treatment?

All generated responses were copied into a separate document without any modification for further analysis.

Readability Assessment: The readability of the generated educational materials was assessed using the following validated readability indices:

- Flesch Reading Ease Score (FRES)
- Flesch–Kincaid Grade Level (FKGL)
- Simple Measure of Gobbledygook (SMOG) Index
- Gunning Fog Index

Readability analysis was performed using standardized online readability assessment tools. Lower grade levels and higher readability scores were considered favorable for patient comprehension.

Accuracy and Quality Evaluation: The scientific accuracy and quality of the generated responses were independently evaluated by three experienced endodontists using a structured evaluation form. The following parameters were assessed:

1. Scientific accuracy
2. Completeness of information
3. Clarity and understandability
4. Presence of misleading information
5. Clinical usefulness for patient education

Each parameter was scored using a 5-point Likert scale:

- 1 = Poor
- 2 = Fair
- 3 = Good
- 4 = Very Good
- 5 = Excellent

The mean scores of all evaluators were calculated for final analysis.

Statistical Analysis: The collected data were entered into Microsoft Excel and analyzed using Statistical Package for Social Sciences (SPSS) software version 25.0. Descriptive statistics including mean, standard deviation, frequency, and percentage were calculated. Inter-examiner reliability among evaluators was assessed using Cohen's kappa coefficient. A p-value of less than 0.05 was considered statistically significant.

Result: The present study evaluated the readability and scientific accuracy of patient education materials generated by ChatGPT for common endodontic conditions and procedures. The readability analysis demonstrated that the generated responses were generally understandable and suitable for individuals with secondary school-level education. The mean Flesch Reading Ease Score was 68.4 ± 5.2 , indicating fairly easy readability, while the Flesch-Kincaid Grade Level, SMOG Index, and Gunning Fog Index suggested moderate reading complexity appropriate for patient education materials (Table 1).

Evaluation of the generated responses by three experienced endodontists revealed good overall scientific accuracy, clarity, and clinical usefulness. Among the evaluated parameters, clarity and understandability showed the highest mean score (4.5 ± 0.4), followed by absence of misleading information (4.4 ± 0.5) and scientific accuracy (4.3 ± 0.5). Completeness of information demonstrated comparatively lower scores, suggesting that certain responses lacked

detailed explanations regarding procedural complications and long-term prognosis (Table 2).

The inter-examiner reliability analysis demonstrated strong agreement among evaluators, with an overall Cohen's kappa value of 0.83, indicating consistency and reliability in the assessment process (Table 3).

Based on the overall quality assessment, the majority of the generated responses were categorized as very good (42%) or excellent (28%), whereas only 8% of the responses were rated as fair. None of the responses were classified as poor quality, indicating the acceptable performance of ChatGPT in generating endodontic educational content (Table 4).

Topic-wise analysis demonstrated higher accuracy scores for general educational topics such as root canal treatment (4.5 ± 0.4) and post-treatment care (4.4 ± 0.5). In contrast, comparatively lower scores were observed for treatment complications (3.7 ± 0.7), suggesting limitations of AI-generated responses in addressing complex clinical scenarios comprehensively (Table 5).

Analysis of the identified limitations revealed that incomplete explanations were the most common issue, observed in 26% of responses, followed by oversimplified information (18%) and lack of emergency guidance (14%). However, 31% of the generated responses demonstrated no major limitations (Table 6).

Overall, the findings of the present study suggest that ChatGPT can generate readable and reasonably accurate endodontic patient education materials; however, expert professional verification remains essential before clinical application.

Table 1. Readability Scores of ChatGPT-Generated Endodontic Patient Education Materials

Readability Parameter	Mean $\pm$ SD	Interpretation
Flesch Reading Ease Score	68.4 ± 5.2	Fairly easy readability
Flesch-Kincaid Grade Level	7.3 ± 1.1	Suitable for grade 7–8 readers
SMOG Index	7.9 ± 1.3	Moderate reading level
Gunning Fog Index	8.1 ± 1.2	Easily understandable by general population

Table 2. Evaluation of Scientific Accuracy and Quality of Information

Parameter Evaluated	Mean Score $\pm$ SD
Scientific Accuracy	4.3 ± 0.5
Completeness of Information	3.9 ± 0.6
Clarity and Understandability	4.5 ± 0.4
Absence of Misleading Information	4.4 ± 0.5
Clinical Usefulness	4.2 ± 0.5

Table 3. Inter-Examiner Reliability for Accuracy Assessment

Parameter	Cohen's Kappa Value	Interpretation
Scientific Accuracy Assessment	0.84	Strong agreement
Completeness Assessment	0.79	Substantial agreement
Clarity Assessment	0.86	Strong agreement
Overall Reliability	0.83	Strong agreement

Table 4. Distribution of Responses Based on Overall Quality Scores

Quality Category	Number of Responses (n)	Percentage (%)
Excellent	28	28%
Very Good	42	42%
Good	22	22%
Fair	8	8%
Poor	0	0%

Table 5. Topic-wise Accuracy Scores of ChatGPT Responses

Endodontic Topic	Mean Accuracy Score $\pm$ SD
Root Canal Treatment	4.5 $\pm$ 0.4
Infected Tooth	4.2 $\pm$ 0.5
Dental Abscess	4.1 $\pm$ 0.6
Post-treatment Care	4.4 $\pm$ 0.5
Treatment Complications	3.7 $\pm$ 0.7
Delayed Treatment Consequences	4.0 $\pm$ 0.6

Table 6. Frequency of Identified Limitations in Generated Responses

Limitation Identified	Frequency (n)	Percentage (%)
Incomplete Explanation	26	26%
Oversimplified Information	18	18%
Lack of Emergency Guidance	14	14%
Missing Long-term Prognosis	11	11%
No Major Limitation	31	31%

Discussion

The rapid advancement of artificial intelligence (AI) in healthcare has significantly influenced the way patients access medical and dental information.¹⁰⁻¹² Conversational AI platforms such as ChatGPT are increasingly being utilized by the general population for obtaining healthcare-related guidance because of their ability to provide quick, simplified, and interactive responses. In dentistry, particularly endodontics, patient education plays a crucial role in reducing anxiety, improving treatment acceptance, and enhancing oral health literacy.^{13,14} Therefore, the present study was conducted to evaluate the readability and scientific accuracy of ChatGPT-generated patient education materials related to common endodontic conditions and procedures.

The findings of the present study demonstrated that the generated educational materials possessed acceptable

readability levels suitable for the general population. The mean Flesch Reading Ease Score observed in the present study indicated fairly easy readability, while the Flesch-Kincaid Grade Level corresponded approximately to secondary school educational standards. These findings are clinically important because patient education materials written at excessively advanced reading levels may reduce patient comprehension and negatively affect treatment-related decision-making. Previous healthcare-based studies evaluating AI-generated educational materials have similarly reported that ChatGPT is capable of producing content with moderate readability suitable for patient communication.

In the present study, the majority of the generated responses demonstrated good scientific accuracy and clarity. The highest scores were observed for clarity and understandability, indicating that ChatGPT effectively simplified endodontic concepts into patient-friendly language. This finding may be attributed to the conversational nature of AI language models, which are specifically designed to generate human-like explanations. Similar observations have been reported in medical literature, where AI-generated educational content was found to be easier for patients to understand compared to conventional scientific text.

Despite the favorable findings, certain limitations in the generated responses were identified. Responses related to treatment complications, emergency management, and long-term prognosis demonstrated comparatively lower completeness scores.^{15,16} Some responses provided oversimplified explanations and lacked detailed clinical information. Although no severely misleading information was identified, incomplete or generalized responses may still create misconceptions if used without professional supervision. This finding highlights the importance of clinician verification before implementing AI-generated educational materials in routine patient care.

The topic-wise analysis revealed that ChatGPT performed better in answering general informational queries such as root canal treatment and post-treatment instructions, whereas comparatively lower accuracy scores were observed for more complex topics such as procedural complications. This may be because AI models are trained to provide generalized educational information and may not consistently interpret nuanced clinical scenarios requiring advanced professional judgment.

The present study also demonstrated strong inter-examiner reliability among evaluators, indicating consistency in the assessment process and strengthening the validity of the findings. Furthermore, a substantial proportion of responses were categorized as very good or excellent quality, suggesting the potential utility of AI-generated educational material as a supplementary communication tool in endodontic practice. The findings of this study should, however, be interpreted in light of certain limitations. The study evaluated responses generated using a limited number of prompts and assessed

only English-language outputs. In addition, the dynamic nature of AI systems means that responses may vary over time depending on software updates and prompt variations. Future research should focus on multilingual evaluation, comparison with other AI platforms, assessment of patient comprehension, and validation of AI-generated content across different dental specialties.

Within the limitations of the present study, ChatGPT demonstrated promising potential in generating readable and reasonably accurate endodontic patient education materials. Nevertheless, professional oversight remains essential to ensure accuracy, completeness, and patient safety before clinical implementation.

Conclusion

Within the limitations of the present study, ChatGPT demonstrated the ability to generate endodontic patient education materials with acceptable readability, good scientific accuracy, and high clarity. Most of the generated responses were understandable for individuals with secondary school-level education and provided clinically useful information regarding common endodontic conditions and procedures. However, certain responses lacked comprehensive explanations related to treatment complications, emergency management, and long-term prognosis. Although no severely misleading information was identified, the possibility of incomplete or oversimplified information highlights the need for professional verification before clinical use.

Therefore, ChatGPT may serve as a useful supplementary tool for patient education in endodontics; however, it should not replace professional dental consultation and expert-reviewed educational resources. Further studies involving larger datasets, multilingual assessment, and patient-based evaluation are recommended to validate its broader clinical applicability.

References

- Alhussain G, Kelly A, O'Flaherty EI, Quinn DP, Flaherty GT. Emerging role of artificial intelligence in global health care. *Health Policy Technol.* 2022 Sep;11(3):100661.
- Jha Kukreja B, Kukreja P. Integration of Artificial Intelligence in Dentistry: A Systematic Review of Educational and Clinical Implications. *Cureus.* 2025 Feb 20;17(2):e79350.
- Casheekar A, Lahiri A, Rath K, Prabhakar KS, Srinivasan K. A contemporary review on chatbots, AI-powered virtual conversational agents, ChatGPT: Applications, open challenges and future research directions. *Comput Sci Rev.* 2024;52:100632.
- Wang L, Wan Z, Ni C, Song Q, Li Y, Clayton E, Malin B, Yin Z. Applications and Concerns of ChatGPT and Other Conversational Large Language Models in Health Care: Systematic Review. *J Med Internet Res.* 2024 Nov 7;26:e22769.
- Funkhouser E, Agee BS, Gordan VV, Rindal DB, Fellows JL, Qvist V, et al. Use of online sources of information by dental practitioners: findings from The Dental Practice-Based Research Network. *J Public Health Dent.* 2014;74(1):71-79.
- Schierz O, Hirsch C, Krey KF, Ganss C, Kämmerer PW, Schlenz MA. Digital dentistry and its impact on oral health-related quality of life. *J Evid Based Dent Pract.* 2024;24(1 Suppl):101946.
- Miller T, Reihlen M. Assessing the impact of patient-involvement healthcare strategies on patients, providers, and the healthcare system: A systematic review. *Patient Educ Couns.* 2023;110:107652.
- Feng S, Li X, Wake AN. Engaging artificial intelligence (AI)-based chatbots in digital health: A systematic review. *PLOS Digit Health.* 2026;5(2):e0001201.
- Singh S, Jamal A, Qureshi F. Readability Metrics in Patient Education: Where Do We Innovate? *Clin Pract.* 2024 Nov 4;14(6):2341-2349.
- Bekbolatova M, Mayer J, Ong CW, Toma M. Transformative Potential of AI in Healthcare: Definitions, Applications, and Navigating the Ethical Landscape and Public Perspectives. *Healthcare (Basel).* 2024 Jan 5;12(2):125.
- Samaranayake L, Tuygunov N, Schwendicke F, Osathanon T, Khurshid Z, Boymuradov SA, Cahyanto A. The Transformative Role of Artificial Intelligence in Dentistry: A Comprehensive Overview. Part 1: Fundamentals of AI, and its Contemporary Applications in Dentistry. *Int Dent J.* 2025 Apr;75(2):383-396.
- Gupta E, Maysan S, Murella S, Saju AR, Grover S, Vasudeva A. Recent dental practices using Artificial Intelligence (AI): A survey. *Bioinformation.* 2025 Mar 31;21(3):514-521.
- Wu X, Zhang B. ChatGPT promotes healthcare: current applications and potential challenges. *Int J Surg.* 2024 Jan 1;110(1):606-608.
- Llorente de Pedro M, Suárez A, Algar J, Díaz-Flores García V, Andreu-Vázquez C, Freire Y. Assessing ChatGPT's reliability in endodontics: Implications for AI-enhanced clinical learning. *Appl Sci.* 2025;15(10):5231.
- Hamada M, Kikuchi S, Akitomo T, Kusaka S, Iwamoto Y, Nomura R. Applications and potential of ChatGPT in dentistry: Scoping review of research perspectives. *J Dent Sci.* 2026;21(1):1-8.
- Achanur M, Bhatt S, Maniyar RN, Sajjanar AK, Roy A, Rao V, et al. ChatGPT's emerging role in dentistry: A review. *J Pharm Bioallied Sci.* 2025;17(Suppl 1):S99-S101