OPEN ACCESS

Saudi Journal of Oral and Dental Research

Abbreviated Key Title: Saudi J Oral Dent Res
ISSN 2518-1300 (Print) | ISSN 2518-1297 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article **Oral Medicine and Radiology**

Efficacy/Accuracy of AI Chatbots to Common Patient Queries for Temporomandibular Joint Disorders: Understandability, Readability, Credibility, Contextual Relevance

Dr. Faisal Taiyebali Zardi[1]*, Dr. Brajesh Gupta[2], Dr. Prabhat Tiwari[3], Dr. Sankojii Supriya[4], Dr Khetavath Prameela[5]

[1]Professor, Department of Oral Medicine and Radiology, Sri Sai College Dental Surgery, Vikarabad, Telangana, India
[2]Professor, Department of Prosthodontics, Sri Sai College of Dental Surgery, Vikarabad, Telangana, India
[3]Professor, Department of Oral and Maxillofacial surgery, Vikarabad, Telangana, India
[4]Post Graduate Student, Department of Oral Medicine and Radiology, Sri Sai College Dental Surgery, Vikarabad, Telangana, India
[5]Assistant Professor, Department of Oral Medicine and Radiology, Malla Reddy Dental College for Women, Hyderabad, Telanagana, India

DOI: https://doi.org/10.36348/sjodr.2026.v11i05.001 | **Received:** 04.03.2026 | **Accepted:** 30.04.2026 | **Published:** 04.05.2026

***Corresponding author:** Dr. Faisal Taiyebali Zardi
Professor, Department of Oral Medicine and Radiology, Sri Sai College Dental Surgery, Vikarabad, Telangana, India

Abstract

Temporomandibular Disorders (TMDs) are a group of conditions affecting the temporomandibular joint and related structures, frequently encountered in dental and medical practice. With the increasing use of large language models (LLMs) for health-related queries, this study aimed to evaluate the quality of responses provided by four leading AI platforms Google Gemini, Microsoft Copilot, ChatGPT, and Meta AI to the most commonly asked questions about TMDs. A set of 45 standardized questions covering definitions, symptoms, causes, diagnosis, and management of TMDs was used. Responses from each platform were assessed on a Likert scale across four key domains: Understandability, Readability, Credibility, and Contextual Relevance. While all four AI models demonstrated potential as educational tools for patients seeking information about TMDs, their quality and consistency varied. Based on this Likert-scale assessment, the overall ranking was: Google Gemini (1st) > Microsoft Copilot (2nd) > ChatGPT (3rd) > Meta AI (4th).

Keywords: TMDs, Google Gemini, ChatGPT, Meta AI, Understandability, Readability, Credibility and Contextual Relevance.

INTRODUCTION

The temporomandibular joint (TMJ) is situated just below the zygomatic arch, in front of the external auditory meatus. It consists of the temporal bone superiorly, the mandible inferiorly, an intra-articular disc, and the masticatory muscles. This joint plays an essential role in enabling jaw movements and is also known as the ginglymoarthrodial joint. [1]

According to the American Academy of Orofacial Pain (AAOP), temporomandibular disorder (TMD) affects the masticatory muscles, the TMJ, and surrounding structures.[2] Common symptoms include pain in the TMJ region, facial or neck discomfort, clicking or popping sounds, joint locking, restricted mouth opening, and difficulty chewing.[3]

Temporomandibular disorders (TMDs) represent one of the most common OROFACIAL problems encountered in clinical practice, manifesting as pain, clicks, muscle fatigue, and functional limitations. Given their chronic nature and impact on quality of life, patients frequently seek additional information online to better understand symptoms, diagnostic approaches, and treatment options.

The growing popularity of large language models (LLMs) has transformed the way patient's access health information. Platforms such as Google Gemini, Microsoft Copilot, Chat GPT, and Meta AI are capable of generating natural language responses to medical queries, positioning themselves as potential educational tools. However, the accuracy, readability, and credibility of such information remain variable.

Although prior studies have explored AI in healthcare, limited literature exists comparing how different LLMs perform specifically in answering TMD-related queries. This study therefore aimed to evaluate and compare the quality of responses generated by four leading AI platforms for commonly asked questions on TMDs.

Gemini is Google's large language model (LLM). More broadly, it's a family of multimodal AI models designed to process multiple modalities or types of data, including audio, images, software code, text and video.

Gemini is also the model that powers Google's generative AI (gen AI) chatbot (formerly Bard) of the same name, much like Anthropic's Claude is named for both the chatbot and the family of LLMs behind it. The Gemini apps on both the web and mobile act as a chatbot interface for the underlying models.[4]

Microsoft Copilot is an artificial intelligence (AI)-powered chat service that integrates with many other Microsoft applications and services to support user efficiency and productivity. Copilot generates content, offers suggestions and automates tasks to support users of various Microsoft technologies, such as Windows, Microsoft 365 and GitHub. Microsoft continues to develop new features and integrations for Copilot.[5]

ChatGPT is an artificial intelligence chatbot from OpenAI that enables users to "converse" with it in a way that mimics natural conversation. As a user, you can ask questions or make requests through prompts, and ChatGPT will respond. The intuitive, easy-to-use, and free tool has already gained popularity as an alternative to traditional search engines and a tool for AI writing, among other things. [5]

The "GPT" in ChatGPT stands for generative pre-trained transformer. In AI, *training* refers to the process of teaching a computer system to recognize patterns and make decisions based on input data, much like how a teacher gives information to their students and then tests their understanding of that information.[6]

ChatGPT is a generative artificial intelligence chatbot developed by OpenAI. It was released in November 2022. It uses large language models specifically generative pre-trained transformers (GPTs) to generate text, speech, and images in response to user prompts.[6]

Meta AI, also known as Meta artificial intelligence, is an AI-driven Chabot powered by LLaMA 3, Meta's newest and most powerful large language model (LLM). Meta introduced an AI technology, LLaMA that can power online chatbots. Meta AI blends into your daily digital interactions by integrating into Meta's chat technology [7].

Meta AI is also the name of the virtual assistant developed by the team, now integrated as a chatbot into Meta's social networking products.[8]

Meta AI is a research division of Meta (formerly Facebook) that develops artificial intelligence and augmented reality technologies.[6]

MATERIALS AND METHODS

A total of 45 standardized questions on TMDs were formulated, covering definitions, etiology, symptoms, diagnostic approaches, and treatment modalities. These questions were posed separately to four AI platforms: Google Gemini, Microsoft Copilot, Chat GPT, and Meta AI.

Each response was assessed on a Likert scale across four key domains by three independent reviewers:

1. **Understandability** – clarity and simplicity for a layperson
2. **Readability** – flow, grammar, and ease of reading
3. **Credibility** – accuracy, evidence-based content, and alignment with standard references
4. **Contextual relevance** – appropriateness and specificity of response to the given question

Scores were tabulated, and compared to determine overall rankings. Statistical analysis was performed using Inter-Rater Agreement (kappa). Statistics revealed a score of 0.80 which is less than p value of 0.05 which makes the results significant.

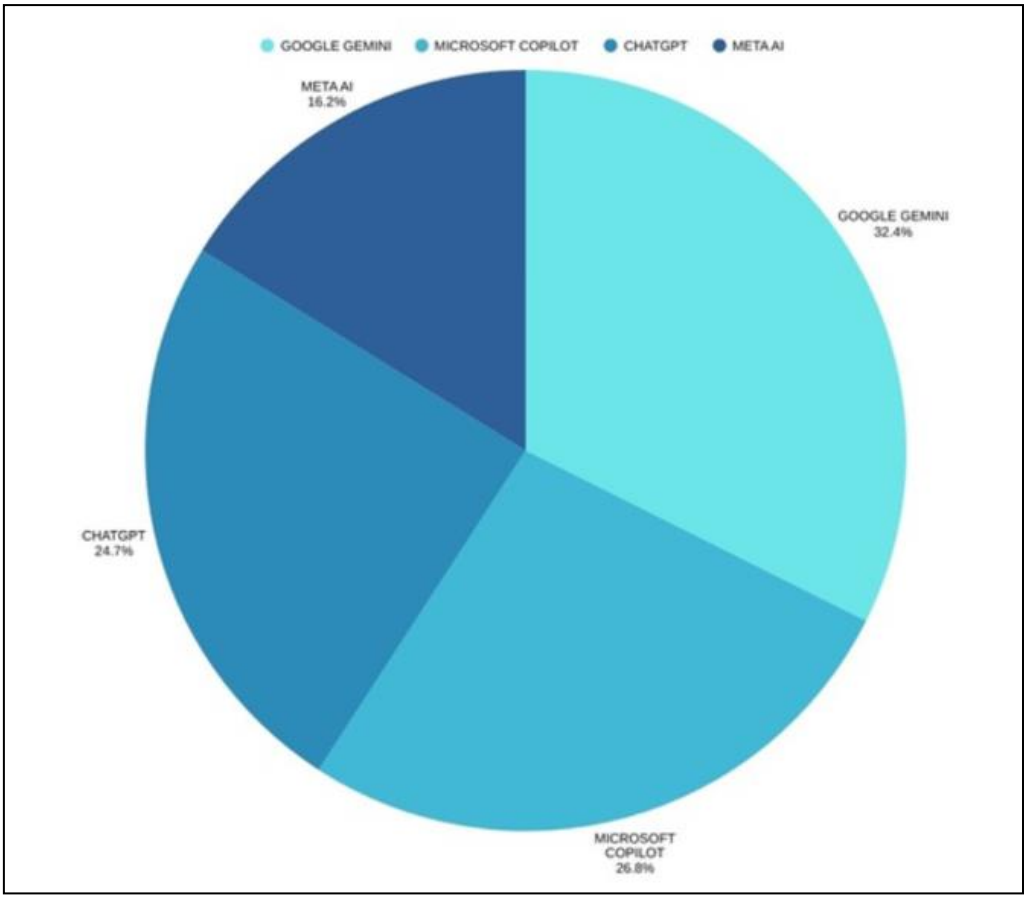

Fig. 1: Pie diagram/chart representing the data of the results

RESULTS

The comparative evaluation revealed significant differences among the four AI systems.

- **Google Gemini** consistently scored highest across all domains, particularly excelling in source referencing and incorporating multimedia elements into its answers.
- **Microsoft Copilot** ranked second, with well-structured explanations.
- **Chat GPT** ranked third; although it provided the broadest and most detailed coverage of all 45 questions.
- **Meta AI** ranked fourth, as its answers are concise.

Overall ranking:

1. Google Gemini 32.4 %
2. Microsoft Copilot 26.8 %
3. Chat GPT 24.7 %
4. Meta AI 16.2 %

DISCUSSION

The results demonstrate that while all four AI platforms are capable of generating useful content on TMDs, their quality varies considerably. Google Gemini's superior readability and credibility suggest it is better suited for patient-facing communication, whereas Chat GPT's extensive detail may appeal more to clinicians or advanced learners. Microsoft Copilot showed balanced performance, providing structured and comprehensible explanations, while Meta AI's limitations highlight the need for further refinement.

The findings are consistent with emerging literature that highlights variability in LLM outputs across healthcare topics. [11] Importantly, reliance on AI-generated health content poses risks, including potential misinformation, oversimplification, or lack of evidence-based guidance. Clinician oversight remains essential to ensure patients receive safe and accurate information.

Another notable observation is that AI-generated responses are generally simplified and user-friendly, enhancing public understanding and awareness. This accessibility may help in early recognition of symptoms and encourage individuals to seek professional care. Nevertheless, there is a potential risk of oversimplification, misinterpretation, or over-reliance on AI tools for self-diagnosis and management, which could delay appropriate clinical intervention.

CONCLUSION

Large language models have promising applications in patient education regarding temporomandibular disorders. However, significant variability exists in their reliability and readability. Based on this study, Google Gemini ranked first,

followed by Microsoft Copilot, Chat GPT, and Meta AI. These findings emphasize the need for cautious integration of AI tools in dental education, with strong clinician oversight to ensure accuracy and safety

AI Platforms demonstrate promising potential in improving public awareness and preliminary understanding of TMD. AI Chatbots may serve as supplementary tools for parental education on TMDs, yet their performance varies by platform. However, their role should be cautiously interpreted, emphasizing the importance of professional consultation for accurate diagnosis and management.

REFERENCES

1. Manchil RD, Chandrasekharan S, Sherly A, Khan F, Lakshmi PS. Evaluation of headache, low back pain, and obstructive sleep apnea, and their relationship in patients with TMJ disorders: A case–control study. J Indian Acad Oral Med Radiol. 2024;36(3):262–267.
2. Franco AL, Fernandes G, Gonçalves DA, Bonafé FS, Camparis CM. Headache associated with temporomandibular disorders among young Brazilian adolescents. Clin J Pain 2014; 30:340–5
3. McNeill C. Management of temporomandibular disorders:Concepts and controversies. J Prosthet Dent 1997; 77:510–22.
4. https://www.ibm.com/think/topics/google-gemini Acceseed on 14-03-2026
5. https://www.techtarget.com/whatis/definition/Microsoft-Copilot Accessed on 10-03-2026
6. https://www.coursera.org/in/articles/chatgpt Accessed on 01-03-2026
7. Meta Platforms, Inc: https://www.softwebsolutions.com Accessed on 10-03-2026
8. Kawale, Ajinkya (20 February 2025). *"India among largest Meta AI adopters, backs open-source innovation"*. Business Standard. Archived from *the original* on 2025-02-20.
9. https://en.wikipedia.org/wiki/Meta_AI Accessed on 12-03-2026
10. https://en.wikipedia.org/wiki/ChatGPT Accessed on 12-03-2026
11. Hassan MG, Abdelaziz AA, Abdelrahman HH, Mohamed MMY, Ellabban MT. Performance of AI-Chatbots to Common Temporomandibular Joint Disorders (TMDs) Patient Queries: Accuracy, Completeness, Reliability and Readability. Orthod Craniofac Res. 2025 Dec;28 Suppl 1: S127-S134.