DOI: https://doi.org/10.54617/adoklinikbilimler.1592592

Original Research Article

Do Artificial Intelligence Chatbots Provide Adequate Information to Patients Using Implant-Supported Prostheses?

Yapay Zeka Sohbet Botları, İmplant Destekli Protez Kullanan Hastalar İçin Yeterli Bilgi Sağlar mı?

Fulya Basmacı[1], Ali Can Bulut[2]

ABSTRACT

Aim: Artificial intelligence (AI) chatbots hold promise with regard to patient education because of their ability to deliver human-like responses to inquiries, yet their reliability in providing accurate information on the use and care of implant-supported prostheses – a critical aspect of prosthodontics – remains uncertain. This study sought to assess the alignment of responses from six AI chatbots to questions on this topic with the current literature on implant-supported prostheses.

Materials and Method: Twenty-five questions related to the usage and maintenance of implant-supported prostheses were posed to six AI chatbots: ChatGPT-4, ChatGPT 01-Preview, ChatGPT 01-Mini, Gemini Advanced, Co-pilot, and Claude 3.5 Sonnet. The accuracy of their responses was assessed by two prosthodontists using a five-point Likert scale, and the average scores were calculated. Differences among the chatbots were analyzed using one-way ANOVA, with the significance level set at α=0.05. As the post-hoc comparison test, Tamhane's T2 test was used.

Results: The accuracy and relevance of the responses provided by the six AI chatbots to questions about the maintenance and use of implant-supported prostheses were evaluated. In terms of accuracy, ChatGPT 01-Preview achieved the highest mean score (4.80±0.08), while Co-pilot received the lowest score (3.22±0.20). ANOVA and Tamhane's T2 tests revealed statistically significant differences between the models ($p<0.05$). Regarding relevance, Claude 3.5 Sonnet obtained the highest mean score (4.94±0.17), whereas Co-pilot demonstrated the worst performance (4.12±0.59).

Conclusion: AI chatbots can serve as effective tools for patient education about implant-supported prostheses. However, inaccuracies in the responses given by certain models and the suboptimal performance of Co-pilot highlight the necessity for human oversight when utilizing these technologies.

Keywords: Artificial intelligence; Chatbot; Dental implant; Digital health; Patient education

ÖZET

Amaç: Yapay zeka (YZ) sohbet robotları, hastaların sorularına insan benzeri yanıtlar vererek hasta eğitiminde kullanılma potansiyeline sahiptir. Ancak protetik diş tedavisinin önemli bir unsuru olan implant destekli protezlerin kullanımı ve bakımı ile ilgili doğru bilgi verme konusundaki güvenilirlikleri ile ilgili bilgiler sınırlıdır. Bu çalışma altı YZ sohbet robotunun implant destekli protezlerle ilgili yanıtlarının güncel literatürle uyumunu değerlendirmeyi amaçlamaktadır.

Gereç ve Yöntem: İmplant destekli protezlerin kullanımı ve bakımıyla ilgili 25 soru, ChatGPT-4, ChatGPT 01-Preview, ChatGPT 01-Mini, Gemini Advanced, Co-pilot ve Claude 3.5 Sonnet olmak üzere altı YZ sohbet robotuna yöneltildi. Yanıtların doğruluğu iki protez uzmanı tarafından beş puanlık Likert ölçeği kullanılarak değerlendirildi ve ortalama puanlar hesaplandı. Sohbet robotları arasındaki farklılıklar tek yönlü ANOVA testi ile analiz edilmiş ve anlamlılık düzeyi α=0.05 olarak belirlenmiştir. Post-hoc karşılaştırma testi olarak Tamhane'nin T2 testi kullanılmıştır.

Bulgular: Altı YZ sohbet robotunun implant destekli protezlerin kullanımı ve bakımıyla ilgili sorulara verdiği yanıtların doğruluğu ve ilgi düzeyi değerlendirildi. Doğruluk açısından ChatGPT 01-Preview en yüksek ortalama puanı (4.80±0.08) alırken, Co-pilot en düşük puanı aldı (3.22±0.20). ANOVA ve Tamhane'nin T2 testleri modeller arasında istatistiksel olarak anlamlı farklar ortaya koydu ($p<0.05$). İlgi düzeyinde ise Claude 3.5 Sonnet en yüksek ortalama puanı (4.94±0.17) alırken, Co-pilot en düşük performansı gösterdi (4.12±0.59).

Sonuç: YZ sohbet robotlarının, implant destekli protezler konusunda verdiği yanıtlar hasta eğitiminde kullanılabileceğini göstermektedir. Ancak, yanıtlardaki bazı yanlışlıklar ve Co-pilot'un yetersiz performansı bu teknolojilerin kullanımında insan denetiminin gerekliliğini ortaya koymaktadır.

Anahtar Kelimeler: Yapay zeka; Sohbet robotu; Dental implant; Dijital sağlık; Hasta eğitimi

Makale gönderiliş tarihi: 28.11.2024; Yayına kabul tarihi: 16.01.2025

İletişim: Dr. Fulya Basmacı
Ankara Yıldırım Beyazıt University Faculty of Dentistry Gazze Cad. No:7 06010 Ayvalı, Etlik, Ankara, Türkiye
E-mail: fulyabasmaci@gmail.com
[1] Asst Prof., Department of Prosthodontics, Faculty of Dentistry, Ankara Yıldırım Beyazıt University, Ankara, Türkiye
[2] Assoc. Prof., Department of Prosthodontics, Faculty of Dentistry, Ankara Yıldırım Beyazıt University, Ankara, Türkiye

INTRODUCTION

Dental implants have become an increasingly frequent treatment option for missing teeth in contemporary dental practice. Despite a decline in the risk of tooth loss in recent years, scientific advancements in implant dentistry have driven the growth of the implant market.[1,2]

The fabrication of an implant-supported restoration involves several critical stages, including ideal planning, surgical placement of the implant, the osseointegration process, and prosthesis construction. Achieving a higher clinical success rate for implants depends not only on the development and refinement of techniques, and materials used for each of these stages, but also on efforts by both patients and professionals to maintain the health of peri-implant tissues.[3,4] Improving patients' knowledge of daily hygiene techniques and potential complications, such as biofilm control and prevention of peri-implant diseases, can significantly impact treatment success.[4-6]

It is essential for patients to receive accurate and effective guidance from their clinicians to understand maintenance instructions for implants and to be motivated to follow them.[4,6] However, even in cases of strong clinician-patient communication, many patients in today's digital era tend to seek additional information from the internet.[4,7]

In this evolving online landscape, artificial intelligence (AI) technologies have become prominent platforms from which patients acquire information. AI is defined as the ability of machines to perform complex tasks such as problem-solving, object and word recognition, and decision-making in a manner that mimics human intelligence. Research on the potential applications of AI in medicine suggests the possibility of these technologies replacing or at least complementing physicians in certain areas.[8] As in other fields of medicine, AI applications have also emerged as promising tools in the field of dentistry.[9] The aim of this study was to analyze the adequacy of responses received by implant-supported prosthesis patients from six different AI programs regarding the maintenance and follow-up of their prostheses and to evaluate the alignment of these responses with the existing literature. The null hypothesis of the study was that the responses provided by the AI programs would align with the literature and would be largely adequate.

MATERIALS AND METHOD

To determine the most frequently searched keywords related to implant-supported restorations, the Google Trends website was utilized. On the basis of this analysis, the search topics "dental implant", "prosthesis", and "implant-supported prostheses" were identified. Subsequently, these keywords were entered into a Google web search using a newly opened browser, and the frequently asked questions listed under the "People also ask" section were reviewed. A total of 25 questions regarding implant-supported prostheses was compiled. The questions primarily focused on cleaning and maintenance of the prostheses, as well as topics related to lifestyle factors such as alcohol and smoking habits, the impact of coffee and sugar, recommended follow-up intervals, and potential complications. These questions were then prepared for evaluation by AI programs (Table 1).

The study involved asking six AI chatbots the 25 questions prepared. The chatbots used were ChatGPT-4 (OpenAI, San Francisco, CA, USA), ChatGPT 01-Preview (OpenAI, San Francisco, CA, USA), ChatGPT 01-Mini (OpenAI, San Francisco, CA, USA), Gemini Advanced (Google LLC, Mountain View, CA, USA), Co-pilot (Microsoft Corp, Redmond, WA, USA), and Claude 3.5 Sonnet (Anthropic, San Francisco, CA, USA).

To minimize any influence from previous responses, each question was posed in a new chat window. The questions were asked in the same order to each AI chatbot, and their responses were recorded. During the evaluation phase, the researchers took additional notes within each session to document their observations and impressions.

Statistical Analysis

The alignment of responses provided by each AI chatbot with the information documented in prosthodontic literature was evaluated by two prosthodontists with academic expertise in the field and experience in training specialists in prosthetic dentistry. The accuracy and relevance of the AI responses were rated using a five-point Likert scale (Table 1).

Table 1. Questions Asked the AI Chatbots about Implant-Supported Prostheses and Likert Scoring

No	Question	A*	R#	LS&
1	Are there different types of implant-supported restorations?			
2	Is hygiene maintenance for implant-supported restorations the same as for natural teeth?			
3	Are hygiene procedures the same for all types of implant-supported restorations?			
4	Is a toothbrush sufficient for cleaning implant-supported restorations? Should I use additional cleaning products?			
5	Should I clean the interproximal areas of implant-supported restorations?			
6	Should I use an oral irrigator for implant-supported restorations?			
7	Should I add cleansing or anti-plaque agents to the oral irrigator?			
8	How should I store my implant-supported prostheses when I remove them at night?			
9	How should I clean the retainers inside my removable implant-supported prostheses?			
10	How often should I attend follow-up appointments for my implant-supported prostheses if I have no issues?			
11	What future issues might I encounter with my implant-supported prostheses?			
12	Under what circumstances should I immediately attend a follow-up if I encounter issues			
13	What should I do to ensure the longevity of my implant-supported prostheses?			
14	Does smoking affect the lifespan of my implant?			
15	Does regular alcohol consumption affect the lifespan of my implant?			
16	Does excessive coffee consumption affect the lifespan of my implant?			
17	Does consuming sugary foods affect the lifespan of my implant?			
18	Do systemic diseases cause differences in the care process of my implants?			
19	Does entering menopause affect my implants?			
20	I've started noticing bad breath – is there a problem with my implant-supported prostheses?			
21	My implant-supported prostheses have started to feel loose – what might be the cause, and what should I do?			
22	The porcelain on my implant-supported prostheses has cracked – what should I do?			
23	My gums bleed when I brush my prostheses – what should I do?			
24	I developed ulcers and sores in my mouth after getting implant-supported prostheses – is this normal?			
25	I keep biting my tongue and cheeks while eating with my implant-supported prostheses – what should I do?			

Likert Scale for Accuracy and Relevance		
Accuracy	**Relevance**	**Likert Score**
Definitely incorrect	Completely irrelevant	1
Incorrect	Irrelevant	2
Partially correct	Partially relevant	3
Correct	Relevant	4
Definitely correct	Completely relevant	5

A*: Accuracy R#: Relevance LS&: Likert Scale

All statistical analyses were conducted using the SPSS for Windows, Version 25.0 (Released 2017, IBM Corp., Armonk, NY, USA). To determine the internal consistency of the researchers' evaluations, the Cohen's kappa value was calculated. The accuracy and relevance scores of the AI models were calculated as mean and standard deviation using descriptive statistics. One-way analysis of variance (ANOVA) was conducted to identify statistical differences among the AI models. To further explore these differences, Tamhane's T2 test was applied for post-hoc comparisons for groups with unequal variances. A statistical significance level of $p<0.05$ was adopted.

RESULTS

Internal Consistency

Cohen's kappa was calculated to assess the inter-rater agreement between the two researchers, yielding a value of $\kappa = 0.76$ ($p<0.001$), which indicates a strong level of agreement. This result supports the reliability of the rating scale and confirms consistency between the two researchers' evaluations. Researcher 1's mean score was 4.50 (SD = 0.82), while Researcher 2's mean score was 4.47 (SD = 0.74). No significant differences were found between the two researchers in terms of mean scores and variability, further demonstrating consistent evaluation standards (Table 2).

Findings Related to the Accuracy of the Responses to the Questions

Descriptive statistics revealed that among the AI models, ChatGPT-01 Preview received the highest accuracy score (4.80±0.08), while Co-pilot had the lowest score (3.22±0.20). ChatGPT-4.0, ChatGPT-01 Mini, and Claude 3.5 Sonn*et al* achieved an accuracy score of 4.78, showing similar performance levels. The Gemini model, with a mean accuracy score of 4.58, displayed moderate performance (Table 3, Figure 1). One-way ANOVA identified statistically significant differences in accuracy scores among the AI models ($p<0.05$) (Table 4). Post hoc analysis using Tamhane's T2 test showed that Co-pilot performed significantly worse than all other models. The largest difference was observed between Co-pilot and Claude, with an average difference of −1.52. Additionally, statistically significant differences were found between ChatGPT-01 Preview and Gemini ($p<0.01$) and between ChatGPT-01 Preview and Claude ($p=0.026$) (Table 4).

Table 2. Inter-Rater Reliability Analysis

Measure	Value	Researcher	Mean Score	Standard Deviation (SD)
Cohen's Kappa (κ)	0.76	Researcher 1	4.50	0.82
Standard Error (SE)	0.05	Researcher 2	4.47	0.74
Significance (p)	<0.001			

Table 3. Mean Accuracy and Relevance Scores for AI Chatbot Responses

Model	Mean Accuracy (±SD)	Mean Relevance (±SD)
ChatGPT-4.0	4.78 (±0.08)	4.88 (±0.31)
Gemini Advanced	4.58 (±0.20)	4.80 (±0.45)
ChatGPT 01-Mini	4.78 (±0.03)	4.88 (±0.37)
Co-pilot	3.22 (±0.20)	4.12 (±0.59)
ChatGPT 01-Preview	4.80 (±0.08)	4.80 (±0.52)
Claude 3.5 Sonnet	4.74 (±0.37)	4.94 (±0.17)

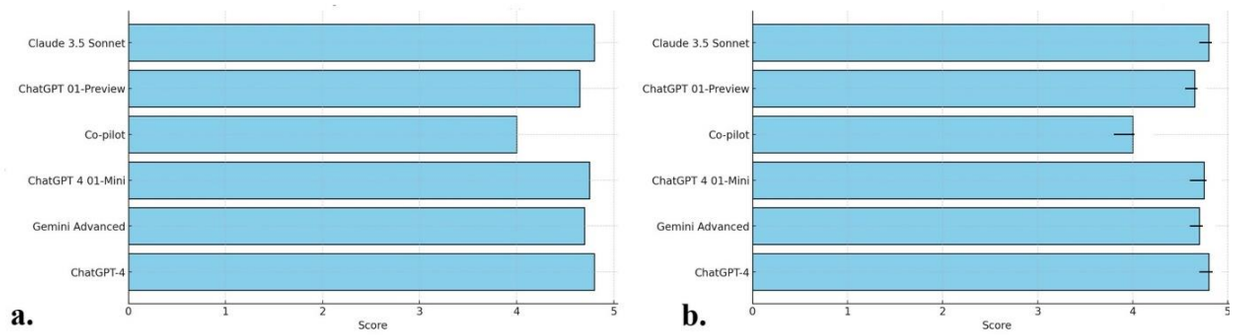

Figure 1. Average Accuracy (a) and Relevance (b) Scores of AI Chatbot Responses to the Questions

Table 4. One-Way ANOVA and Post Hoc Tamhane's T2 Test Results for Accuracy of AI Chatbot Responses

Researcher	Source	SS	df	MS	F	P Value	Model Pair	Mean Differences (I-J)	Std Error	P Value	95% CI (Lower)	95% CI (Upper)
1	Between Groups	62.14	5	12.42	47.90	<0.001*	Co-pilot - Claude	-1.52	0.12	0.001*	-1.87	-1.17
	Within Groups	37.36	144	0.26			Co-pilot - ChatGPT-4.0	-1.56	0.14	0.001*	-1.99	-1.13
2	Between Groups	41.41	5	8.28	29.88	< 0.001*	Co-pilot - Gemini	-1.36	0.14	0.001*	-1.80	-0.92
	Within Groups	39.92	144	0.28			Co-pilot - 01-Mini	-1.56	0.15	0.001*	-2.03	-1.09
							Co-pilot - 01Preview	-1.62	0.15	0.001*	-1.83	-1. 41
							01Preview - Gemini	0.48	0.12	0.003*	0.12	0.84
							01Preview - Claude	0.44	0.13	0.026*	0.03	0.85

Statistically significant differences are indicated with an asterisk (*) ($p < 0.05$)

Findings Related to the Relevance of Responses to Questions

An analysis of the descriptive statistics for the relevance of the responses showed that the Claude 3.5 Sonnet model achieved the highest performance with a mean score of 4.94 ±0.17. In contrast, Co-pilot was identified as the lowest-performing model with a mean score of 4.12 ±0.59. The scores for the other models ranged between 4.80 and 4.88 (Table 3, ure 1). ANOVA revealed statistically significant differences in relevance scores among the AI models assessed by Researcher 1 and Researcher 2 ($p<0.05$) (Table 5). Post-hoc analysis using Tamhane's T2 test showed that Co-pilot performed significantly worse than the other models ($p<0.05$). The largest difference was observed between Co-pilot and Claude, with a mean difference of −0.82. No significant differences were found between ChatGPT-4.0, Gemini, and ChatGPT-01 Mini. Similarly, no significant difference was observed between ChatGPT-01 Preview and Claude 3.5 Sonnet ($p=0.48$) (Table 5).

Table 5. One-Way ANOVA and Post Hoc Tamhane's T2 Test Results for Relevance of AI Chatbot Responses

Researcher	Source	SS	df	MS	F	P Value	Model Pair	Mean Differences (I-J)	Std Error	P Value	95% CI (Lower)	95% CI (Upper)
1	Between Groups	14.54	5	2.91	13.53	<0.001*	Co-pilot - Claude	-0.82	0.12	0.001*	-1.21	-0.43
	Within Groups	30.96	144	0.22			Microsoft Go ChatGPT-4.0	-0.78	0.14	0.001*	-1.19	-0.37
2	Between Groups	10.22	5	2.04	10.70	<0.001*	Co-pilot - Gemini	-0.68	0.16	0.001*	-1.16	-0.20
	Within Groups	27.52	144	0.19			Microsoft Go 01-Mini	-0.76	0.14	0.001*	-1.23	-0.29
							Co-pilot - 01Preview	-0.68	0.16	0.004*	-1.17	-0.19
							01Preview - Claude	0.00	0.15	0.48	-0.49	0.49
							01Preview - Gemini	0.00	0.18	1.0	-0.64	0.64

Statistically significant differences are indicated with an asterisk (*) ($p < 0.05$)

Researcher's Observations Regarding the AI models

Gemini Advanced: Strengths: Demonstrated excellence in providing detailed explanations and utilizing professional language. Offered noteworthy suggestions, such as recommending mouthguards during sports to protect the prosthesis and addressing specific habits (e.g., nail biting, chewing on pens). Weaknesses: Responses were frequently too general, lacking the specificity required to address more targeted questions effectively.

ChatGPT 01-Mini: Strengths: Excelled in providing detailed information on systemic conditions, such as the effects of autoimmune diseases and medications on prostheses. Offered practical tips for temporary measures, which was seen as a positive feature. Weaknesses: Some responses were overly technical (e.g., use of terms like "removable implant prostheses") or irrelevant (e.g., temporary solutions for porcelain cracks).

ChatGPT 01-Preview: Strengths: Presented answers in a clear cause-effect framework, enhancing understandability. Demonstrated a strong ability to offer additional practical suggestions, such as recommending night guards for bruxism and teeth grinding. Weaknesses: Responses to questions regarding substance abuse (e.g., alcohol and smoking) were superficial and lacked sufficient depth or detail.

Claude 3.5 Sonnet: Strengths: Provided detailed recommendations for implant care, such as suggesting non-abrasive toothpaste and specific brushing techniques. The only model to address the effects of electronic cigarettes in relation to smoking. Weaknesses: Some responses could have been more detailed in addressing specific questions.

Microsoft Co-Pilot: Strengths: Supported its information with references, which added credibility. Weaknesses: Responses were brief and superficial, leading to overall lower performance compared to other models.

ChatGPT-4: Strengths: Delivered high-quality responses with sufficient information for most questions. Weaknesses: Responses were less detailed compared to Claude 3.5 Sonnet.

General Observations Across All Models

Toothbrush and Toothpaste Recommendations: All models recommended soft-bristle toothbrushes and fluoride toothpaste. However, further scientific validation of these recommendations is necessary. Salt-

water Rinse: Frequently suggested by the models. While this aligns with literature, the scientific rationale for its use was not consistently explained. Prosthesis Cleaning: Daily cleaning with denture cleansers was recommended, but the abrasive effects of such products were overlooked. Follow-up Schedule: Some models suggested monthly follow-ups during the first three months. This advice, while somewhat aligned with guidelines, requires confirmation from a clinician. Response Time and Level of Detail: ChatGPT 01-Mini provided the longest and most detailed responses. Microsoft Co-Pilot delivered the shortest and most superficial answers.

DISCUSSION

Dental implant treatments are among the most frequently applied procedures in prosthodontic clinical practice today. The long-term success of implant treatments depends not only on an ideal surgical procedure and a prosthesis fabricated under optimal conditions but also on a personalized follow-up and maintenance program. In this context, the dentist-patient relationship and ensuring that patients are well-informed play crucial roles in maintaining patient motivation.

Beyond their clinicians, patients often turn to the internet for additional information about their dental implants. However, the reliability of information available online can vary significantly, and such sources may contain incomplete or inaccurate information.[7,10] These tools have the potential to provide standardized, accessible, and consistent answers, but their efficacy in delivering accurate and detailed guidance must be rigorously evaluated, particularly in critical areas like implant maintenance and care.

This study analyzed the adequacy of responses provided by commonly used AI chatbots to questions regarding the maintenance and follow-up of implant-supported prostheses and evaluated the alignment of these responses with the literature. The null hypothesis, which posited that the responses generated by AI chatbots would align with the literature and be largely adequate, was accepted.

The findings underscore the potential of AI chatbots as supplementary tools for patient education in prosthodontics. However, the variability in response quality among the different AI models highlights the need for critical assessment and oversight when relying on such technologies in clinical and educational contexts.

The results of this study are consistent with previous research in the literature that has explored the use of AI chatbots for patient education in various fields of dentistry. For instance, Yurdakurban *et al*.[11] evaluated the data quality of AI chatbots in informing patients undergoing orthognathic surgery. Their findings indicated that chatbots generally provided high-quality and reliable answers to patient inquiries. Similarly, Polizzi *et al*.[12] analyzed the potential of AI for personalized treatment planning in patients with periodontitis through a systematic review. Their study highlighted the potential of AI algorithms to improve accuracy and reliability in predicting future periodontitis. These studies, along with the current findings, emphasize the promise that AI will help enhance patient education and personalized care within dentistry. However, the need for further refinement in AI responses and validation against clinical standards remains evident. Jacobs *et al*.[13] evaluated whether AI could serve as a patient-friendly and accurate resource for third molar extraction. Comparing AI responses with the guidelines of the American Association of Oral and Maxillofacial Surgeons, they found that while there were minor inaccuracies or omissions, most responses were accurate. Similarly, Dursun *et al*.[14] assessed the effectiveness of AI models as patient advisors in orthodontics. They reported that while AI chatbots generally provided correct and moderately reliable answers, the readability of the responses posed challenges for patients. In a randomized parallel-group study conducted with 224 patients in India, Ghosh *et al*.[15] utilized a semi-autonomous AI system for patient recall and observed an increase in recall rates from 21.1% to 37.8%. AI chatbots could potentially improve follow-up attendance and oral hygiene habits among implant-supported prosthesis patients in a similar way, contributing to increased survival rates for prosthetic treatments.

The success of implant restorations depends on appropriate indications and planning, ideal surgical and prosthetic procedures, improvement of the patient's oral hygiene, and the management of potential complications. During the delivery session, patients are typically informed by their clinicians or dental hygienists about the guidelines for prosthesis maintenance.

Battista *et al*.[16] examined ChatGPT's ability to generate consent forms for surgical risks associated with dental implant placement in patients with periodontal issues, smoking habits, or diabetes. Their study found that AI-generated consent forms performed as well as those written by humans and were significantly better for diabetic patients. They concluded that ChatGPT has the capability to independently produce accurate and useful patient education and management documents.

These findings collectively highlight the potential of AI chatbots to improve patient outcomes by providing accurate, accessible, and individualized information. However, challenges such as minor inaccuracies, readability issues, and the need for professional oversight underscore the importance of integrating AI into clinical practice with caution and thoroughly validating such use.

Another critical factor influencing the success of implant-supported restorations is the dentist's ability to identify the individual and specific personal care needs of each dental implant patient. The recommendations provided to patients must be tailored according to various parameters, including prosthesis design, the position and angulation of implants, the length and placement of transmucosal abutments, smoking habits, oral hygiene practices, and manual dexterity.

To maintain optimal peri-implant health, patients must adhere to a consistent regimen of daily oral care and periodic professional maintenance. Tailoring these recommendations to each patient's unique circumstances ensures not only the longevity of the prosthetic treatment but also minimizes the risk of complications, ultimately enhancing the overall success of implant-supported restorations.

Lyle *et al*.[17] recommend oral irrigation devices as a method to aid in biofilm control in hard-to-reach areas around implants and implant-supported prostheses. Kracher *et al*.[18] further noted that while irrigation devices are effective in removing biofilm and food debris, improper use or excessive water pressure can damage the junctional epithelium, and these devices should thus be used at low pressure. Researchers generally advise a variety of tools for daily care of implants and peri-implant tissues, including manual or electric toothbrushes, various types of dental floss, interdental brushes, mouth rinses, and oral irrigation devices.[19] In the present study, the AI chatbots provided similar responses to questions two to seven regarding these recommendations.

The responses of the chatbots to questions about smoking and implant-supported restorations also aligned with the literature. The impact of smoking on the risk of periodontal disease has been well-documented for years. Although the risks are dose-dependent, many professionals argue that smoking should be considered an absolute exclusion criterion for implant placement and recommend smoking cessation prior to the procedure.[20]

Ferro *et al*.[4] emphasized that AI is an effective tool for increasing patient motivation and demonstrated that interactive technologies positively influence patient compliance with treatment. They observed that AI played a strong role in highlighting the negative effects of smoking on peri-implant health. Similarly, Banerjee and Shehab's findings showed that AI not only enhances the educational process but also improves patient adherence and clinical outcomes. For example, increasing awareness of the peri-implant disease risks associated with behaviors like smoking led to improved treatment success. Furthermore, regular reminders about the importance of maintenance and follow-ups positively impacted the long-term success of implants.[20,21] In this study, all the AI chatbots recommended smoking cessation to patients, emphasizing its benefits for both the health of implant-supported restorations and general health.

Patient education and understanding of maintenance protocols are crucial for the success of dental implant treatments. Traditional methods, often limited to one-on-one education and written instructions from clinicians, can be significantly enriched with the integration of AI-based systems. For instance, Banerjee *et al*.[21] highlighted the supportive role of AI in patient education and maintenance processes, reporting that chatbots like ChatGPT provided highly accurate and contextually relevant responses, particularly in preventing peri-implant diseases. Clinical recommendations such as the use of antibacterial mouth rinses, soft-bristle toothbrushes, and smoking cessation were effectively communicated.

AI's potential for personalized patient education has also been emphasized in various other studies.[4,21] By considering factors such as a patient's age, overall health status, type of implant used, and condition of peri-implant tissues, AI algorithms can tailor their suggestions. This allows patients to better understand and implement care instructions that are specific to their needs. Personalized guidance holds greater value than generic advice, improving compliance and outcomes.

The present study found that certain AI models excelled in providing accurate and relevant information tailored to individual needs. These results underline the growing potential of AI chatbots as tools for enhancing patient education and ensuring the success of implant-supported prostheses. Shehab *et al*.[22] explored the potential of AI in improving health literacy, emphasizing the ability of AI-based systems to simplify complex medical information into a more patient-friendly format. This capability is particularly beneficial for individuals with low health literacy, enabling them to better understand and implement post-implant care protocols. For instance, simple yet effective instructions for reducing peri-implant infection risks or maintaining the hygiene of tissues surrounding implants can be made more comprehensible through AI-generated guidance. Shehab's study also highlighted the role of AI in overcoming language barriers. Multilingual AI systems can serve global patient populations more effectively, bridging communication gaps and enhancing inclusivity in healthcare delivery. This observation aligns with the feedback from researchers in this study, particularly regarding the overly technical language used in some of ChatGPT 01-Mini's responses.

Regarding post-delivery care for implant-supported prostheses, it is recommended that radiographic evaluations be performed one year after delivery and subsequently every two years.[15,23] In this study, the AI chatbots suggested follow-up appointments starting at three months, with Claude 3.5 Sonnet, ChatGPT 01-Preview, ChatGPT-4, and ChatGPT 01-Mini providing the most accurate information in this regard.

The impact of systemic diseases on implant-supported restorations is well-documented,[1,2,18,19] and all AI chatbots in this study addressed this topic effectively, demonstrating their ability to incorporate clinically relevant systemic factors into their responses.

Considering all these parameters, it is essential to recognize that the maintenance and follow-up processes for implant-supported prostheses are nevertheless dynamic and should be tailored to the individual patient based on their unique risks and needs. While AI chatbots show promise as a basic information source, the accuracy and currency of their responses must be verified. Moreover, their lack of contextual understanding and emotional support capabilities limits their utility, underscoring the necessity of professional oversight and patient-specific protocols developed by clinicians.

As AI technology continues to evolve, ongoing research and development aimed at addressing the shortcomings of chatbots – such as by enhancing contextual comprehension and emotional engagement – will make them more reliable and effective tools for patient education. Additionally, AI's potential to function as an equalizer in healthcare by providing accessible and standardized information to patients with varying levels of access to care is noteworthy.

The limitations of this study include the ever-evolving nature of chatbots, which may lead to variations in their responses, and the narrow scope of the questions posed, which represent only a subset of the questions patients might ask.

CONCLUSION

AI chatbots demonstrated strong performance in educating implant-supported prosthesis patients by providing accurate responses, highlighting their significant potential in improving health literacy and patient education, particularly for those with limited access to healthcare services. However, patients using AI-based technologies should verify the accuracy of the information provided. Further research is needed to enhance the reliability of these systems through simulated scenarios and personalized instructions.

CONFLICT OF INTEREST STATEMENT

No conflict of interest was declared by the authors.

REFERENCES

1. Alghamdi HS, Jansen JA. The development and future of dental implants. Dent Mater J 2020;39:167-72.

2. Buser D, Sennerby L, De Bruyn H. Modern implant dentistry based on osseointegration: 50 years of progress, current trends and open questions. Periodontol 2000 2017;73:7-21.

3. Monje A, Aranda L, Diaz KT, Alarcón MA, Bagramian RA, Wang HL, *et al*. Impact of maintenance Therapy for the prevention of peri-implant diseases: A systematic review and meta-analysis. J Dent Res 2016;95:372-9.

4. Ferro AS, Nicholson K, Koka S. Innovative trends in implant dentistry training and education: a narrative review. J Clin Med 2019;8:16-8.

5. Pulcini A, Bollaín J, Sanz-Sánchez I, Figuero E, Alonso B, Sanz M, *et al*. Clinical effects of the adjunctive use of a 0.03% chlorhexidine and 0.05% cetylpyridinium chloride mouth rinse in the management of peri-implant diseases: A randomized clinical trial. J Clin Periodontol 2019; 46:342-53.

6. Walia K, Belludi SA, Kulkarni P, Darak P, Swamy S. A comparative and a qualitative analysis of patient's motivations, expectations and satisfaction with dental implants. J Clin Diagn Res 2016;10:23-6.

7. Menziletoglu D, Guler AY, Isik BK. Are YouTube videos related to dental implant useful for patient education? J Stomatol Oral Maxillofac Surg 2020;121:661-4.

8. Ahuja AS. The impact of artificial intelligence in medicine on the future role of the physician. Peer J 2019; 4;7:e7702.

9. Reyes TR, Knorst JK, Ortiz FR, Ardenghi TM. Scope and challenges of machine learning-based diagnosis and prognosis in clinical dentistry: A literature review. J of Clin Transl Res 2021;7:523-39.

10. Berge TI. Public awareness, information sources and evaluation of oral implant treatment in Norway. Clin Oral Implants Res 2000;11:401-8.

11. Yurdakurban E, Topsakal KG, Duran GS. A comparative analysis of AI-based chatbots: Assessing data quality in orthognathic surgery related patient information. J Stomatol Oral Maxillofac Surg 2024;125:101757.

12. Polizzi A, Quinzi V, Lo Giudice A, Marzo G, Leonardi R, Isola G. Accuracy of artificial intelligence models in the prediction of periodontitis: A systematic review. JDR Clin Trans Res 2024;9:312-24.

13. Jacobs T, Shaari A, Gazonas CB, Ziccardi VB. Is ChatGPT an Accurate and readable patient aid for third molar extractions? J Oral Maxillofac Surg 2024;82:1239-45.

14. Dursun D, Bilici Geçer R. Can artificial intelligence models serve as patient information consultants in orthodontics? BMC Med Inform Decis Mak 2024;24:211.

15. Ghosh A, Saha AP, Saha S, Das A. Promoting the importance of recall visits among dental patients in India using a semi-autonomous AI system. Stud Health Technol Inform 2022;16:85-92.

16. Di Battista M, Kernitsky J, Dibart S. Artificial intelligence chatbots in patient communication: current possibilities. Int J Periodontics Restorative Dent 2024;44:731-8.

17. Lyle DM. Implant maintenance: is there an ideal approach? Compend Contin Educ Dent 2013;34:386-90.

18. Kracher CM, Smith WS. Oral health maintenance dental implants. Dent Assist 2010;79:27-35.

19. Louropoulou A, Slot DE, Van der Weijden F. Mechanical self-performed oral hygiene of implant supported restorations: a systematic review. J Evid Based Dent Pract 2014;14:60-9.

20. Carra MC, Blanc-Sylvestre N, Courtet A, Bouchard P. Primordial and primary prevention of peri-implant diseases: A systematic review and meta-analysis. J Clin Periodontol 2023;50:77-112.

21. Banerjee, TN, Paul P, Debnath A, Banerjee S. Unveiling the prospects and challenges of artificial intelligence in implant dentistry. A systematic review. J Osseointegration 2024; 16: 53-60.

22. Shehab AA, Shedd KE, Alamah W, Mardini S, Bite U, Gibreel W. Bridging gaps in health literacy for cleft lip and palate: The role of artificial intelligence and interactive educational materials. The Cleft Palate Craniofacial J 2024;9:1-7.

23. Reading K, Knowles L, Towns S. The hygienist's role in the management of the implant patient in primary care. Prim Dent J 2024;3:53-62.