e-ISSN: 2980-1389

JDSE

Journal of Dental Sciences and Education

MHA MEDIHEALTH ACADEMY

Journal of Dental Sciences and Education

CONTENTS

Volume: 4 Issue: 3 Year: 2026

ORIGINAL ARTICLES

REVIEW

CASE REPORT

Journal of Dental Sciences and Education

Original Article

DOI: 10.51271/JDSE-0081

A multidisciplinary evaluation of AI-powered chatbots on apical root-end resection-assessing alignment with international endodontic guidelines: a comparative methodological study

Tuğçe Paksoy*[1], Selin Gaş[2], Seval Ceylan Şen[3], Özlem Saraç Atagün[3], Gülbahar Ustaoğlu[3], Şeyma Çardakçı Bahar[3], Savaş Özarslantürk[4]

[1]Department of Periodontology, Hamidiye Faculty of Dentistry, University of Health Sciences, İstanbul, Turkiye
[2]Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, İstanbul Gelişim University, İstanbul, Turkiye
[3]Department of Periodontology, Gülhane Faculty of Dentistry, University of Health Sciences, Ankara, Turkiye
[4]Department of Oral and Maxillofacial Radiology, Gülhane Faculty of Dentistry, University of Health Sciences, Ankara, Turkiye

Cite this article as: Paksoy T, Gaş S, Ceylan Şen S, et al. A multidisciplinary evaluation of AI-powered chatbots on apical root-end resection-assessing alignment with international endodontic guidelines: a comparative methodological study. *J Dent Sci Educ.* 2026;4(3):76-83. doi:10.51271/JDSE-0081

Received: 22.05.2026 ◆ **Accepted**: 30.07.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Aims: This study aimed to identify the clinically relevant patient questions about apical root-end resection based on international consensus guidelines, and to systematically evaluate the accuracy, quality, usefulness, and readability of generated responses by four Artificial Intelligence (AI)-powered conversational agents, ChatGPT-4, DeepSeek V3.1, Gemini 3, and Copilot.

Methods: A standardized set of 20 clinical questions was developed from the American Association of Endodontists guidelines and international consensus reports. A priori, an evidence-based answer key was established as the reference standard. Queries were submitted to ChatGPT-4, DeepSeek V3.1, Gemini 3, and Copilot, following a standardized interaction protocol designed to minimize personalization and carryover effects, such as incognito mode, newly created accounts, and no response regeneration. The outputs were independently assessed by three experts in periodontics and oral radiology, using validated tools, such as CLEAR criteria, modified Global Quality Score (mGQS), accuracy ratings, DISCERN, readability indices (Flesch reading ease, FRE and Flesch-Kincaid grade level, FKGL) and the patient education materials assessment tool (PEMAT).

Results: ChatGPT-4 and Copilot achieved significantly higher CLEAR scores than Gemini 3. Copilot achieved significantly higher mGQS and accuracy than DeepSeek V3.1 and Gemini 3 and received significantly lower, and therefore more favorable, usefulness scores. Gemini 3 scored higher on the FKGL test and lower on the understandability test compared to DeepSeek V3.1. Relationships between evaluation metrics were different between models.

Conclusion: The AI models evaluated demonstrated significant variability in clinical accuracy, quality, readability, and patient-centered usability. Copilot had the best overall mGQS and Accuracy scores and the best Usefulness scores, while the models varied in their strengths in readability, understandability, and actionability.

Keywords: Artificial Intelligence, chatbots, ChatGPT, patient education, readability

INTRODUCTION

Microbial infection is a key factor in the development of pulpal and periapical diseases.[1] The main goals of endodontic therapy are to eliminate or greatly reduce the number of microorganisms within the root canal system and to place a proper seal to prevent re-infection of the periapical tissues.[2] Conventional root canal therapy has an approximate 90% success rate, but failure may require nonsurgical retreatment. Where this is not possible, surgical endodontic therapy is indicated.[3] Apical root-end resection is mainly indicated in the surgical treatment of teeth with advanced endodontic-periodontal involvement among surgical approaches, but it is also indicated in some special cases of surgical endodontic management as described by the American Association of Endodontists (AAE) and Cohen's Pathways of the Pulp.[4,5]

Patients are increasingly supplementing the professional advice they receive from dentists with digital sources. Patients want to know the risks of procedures, how long recovery will take, and what the expected results are. The transition to Artificial Intelligence (AI)-powered platforms emphasizes the importance of delivering not just accurate but also patient-centered and accessible information.[6,7] The most complex part of the internal root canal anatomy is the apical root canal anatomy, in which irregularities such as isthmuses and accessory canals are often present.[8,9] These anatomical variations of the apical third may pose a challenge

***Corresponding Author:** Tuğçe Paksoy, tugce.paksoy@sbu.edu.tr

for endodontic procedures and are considered important risk factors for treatment failure.[10,11] When orthograde root canal therapy fails, apical root-end resection often is advocated to remove persistent microbial reservoirs in the complex apical anatomy. It has been reported that removal of the apical 3 mm of the root removes approximately 92% of such irregularities.[10] However, apical root-end resection may lead to thermal changes and procedural alterations in the root dentin, which may harm the periodontal ligament cells and adjacent alveolar bone.[12] Thus, careful control of temperature elevation and minimally invasive surgical techniques during apical root-end resection are essential to minimize the risk of damage to surrounding vital soft and hard tissues. Therefore, patients may seek information about indications, risks, success rates, and postoperative course in online resources such as AI chatbots.

Since its birth in 1956, AI, loosely defined as computer systems that mimic human thought processes, has become a revolutionary technology field.[13] While AI has the potential to revolutionize clinical dentistry, a major concern for patient safety is the risk of 'hallucinations', where models generate information that sounds convincing but is clinically wrong or out of date. Hence, a thorough assessment of AI-produced content is highly desirable to guarantee the correctness and understandability of information about complicated surgical interventions like apical root-end resection.[14,15] AI, which is the ability to accomplish tasks normally associated with human intelligence including learning, reasoning, problem-solving and decision-making, has made great strides in recent years due to improvements in computational power, algorithmic structures and big data analysis. These advances have paved the way for widespread use of AI in many areas, including the health sciences. The use of AI is gaining more acknowledgement in dentistry as a way to increase diagnostic accuracy, improve treatment planning, and advance education.[6,7,16]

While the performance of AI has been investigated in many dental fields including prosthetics, hypersensitivity and implantology, there are relatively fewer studies concerning the reliability of AI agents in endodontic surgery.[16-18] The present study provides a multidimensional evaluation of four prominent AI models to bridge this gap. A key component of this work is to assess the language difficulty of the AI outputs, as health information materials should be written at or below an eighth-grade reading level to optimize public understanding.[19,20] With the increasing use of AI chatbots by patients to seek oral health information, the objective of this study is to assess the clinical utility of AI chatbots as adjunctive educational tools in current endodontic practice using a comprehensive set of validated tools, namely, CLEAR, mGQS, DISCERN, and PEMAT.

METHODS

Ethics

No ethics committee approval was necessary, as the study did not involve human participants, identifiable patient data, or biological materials. All procedures were carried out in accordance with the ethical rules and the principles.

Study Design and Question Development

The standard question set was developed using authoritative clinical and anatomical references, including international guidelines, consensus-based references and relevant peer-reviewed studies on surgical endodontics. The core clinical themes and patient concerns were based on the AAE Glossary and Clinical Guidelines and Cohen's Pathways of the Pulp[4,5,10] and well-established anatomic success criteria, which demonstrate that a 3 mm apical root-end resection removes around 92% of the apical ramifications. Finally, a set of 20 questions was developed to cover clinical indications, success rates, procedural risks and postoperative protocols. Two authors (TP and SCŞ, periodontology specialists) reviewed the items for linguistic and clinical accuracy and they eliminated duplicates and semantically equivalent questions and refined the items to ensure grammatical clarity.

AI Interaction and Data Collection

A primary account was created for each of the four platforms (ChatGPT-4, Gemini 3, DeepSeek V3.1 and Copilot) with unique credentials and accessed in incognito mode to remove any interference from previous user data or search history. Each query was posed as a 'new chat' by a member of the research team (SG, specialist in oral and maxillofacial surgery) to ensure independence between responses. Only the first output was collected and the regeneration feature was avoided strictly for the sake of consistency of the data. All interactions occurred on 20th November, 2025 between 09:00 and 16:00.

Reference Standard and Outcome Assessment

To facilitate an objective evaluation of the AI-generated responses, a standardized "answer key" (reference standard) was created prior to the evaluation phase grounded on clinical consensus statements.[4,5,10] The answers were evaluated by three independent specialists, who were unaware of the question formulation team: OSA and GUS (periodontology specialists) and SÖ (specialist in oral and maxillofacial radiology). The multidisciplinary panel graded the AI agents using the evidence-based reference key and the CLEAR criteria, modified Global Quality Score (mGQS), accuracy ratings, the DISCERN instrument, standard readability indices (Flesch Reading Ease, FRE and Flesch-Kincaid Grade Level, FKGL), and the patient education materials assessment tool (PEMAT). To reduce subjective inter-observer variability and guarantee a thorough comparison with existing clinical evidence, the validation and data curation process was overseen by another periodontology expert (ŞÇB).

To examine whether the repeated values of the same variables are similar, intra-class correlation (two-way mixed and consistency approach) was used. The analyses showed that the inter-rater reliability level was above the minimum value of 0.700 and the inter-rater reliabilities were statistically significant and highly consistent ($p<0.05$).

The qualitative integrity of the AI outputs was assessed with the CLEAR tool, which divides reliability into five core domains: content completeness, no misinformation, evidence-based grounding, appropriate formatting, and topical relevance. The evaluators scored each category (1-5) and a total quality index was calculated; higher scores indicated better reliability of the information.[21]

The mGQS was used to assess overall contextual depth and accuracy of the answers. This 5-point scale ranges from 'strongly disagree' (indicating entirely flawed content) to 'strongly agree' (signifying a comprehensive and error-free response).[22]

Accuracy was measured using a five-point Likert scale. The scale is as follows; Score 1: responses are completely wrong; Score 2: responses contain more wrong than right information; Score 3: responses contain an equal amount of wrong and right information; Score 4: responses contain more right than wrong information; and Score 5: responses are completely right.[23]

The usefulness of the AI-generated responses was assessed using a structured 4-point rating scale. This scale characterizes the responses as follows: (score 1) very useful: the responses are complete and correct, like what a domain expert would provide. (Score 2) useful; the responses are mostly correct but may miss some important details. (Score 3) Partially useful: the responses contain both correct and incorrect/misleading information. (Score 4) Not useful: the responses contain only incorrect/misleading information, with no reliable content.[24] In this study, usefulness was defined as the extent to which a response provided practical information that was needed by the patient, and provided relevant, understandable and clinically applicable guidance. Lower scores indicated more usefulness. A negative correlation with positively oriented measures like CLEAR, mGQS and Accuracy was thus due to an opposite direction of scoring scales, and not to an opposite conceptual relationship.

The linguistic accessibility was quantified using the FRE and FKGL metrics. The FRE is a numerical representation of the difficulty of a text (100 being the most easily understood) and the FKGL transforms this into the American educational grade level required for complete comprehension.

To systematically evaluate these two key aspects of patient education materials, a standardized and validated tool called the PEMAT was employed to assess the comprehensibility and actionability of the responses generated by AI. The generated advice was tested for understandability by lay people and actionability in a clinical setting using PEMAT.[19,20]

Individuals with different cultural and educational backgrounds and varying levels of health literacy must be able to understand and correctly interpret the main messages of instructional materials for them to be considered intelligible. Understandability refers to the extent to which individuals with varying levels of health literacy can comprehend the principal messages of educational material. Actionability refers to the degree to which users are able to identify specific actions that can be taken based on the information. The understandability domain of the PEMAT contains 17 items that are rated 1 (agree), 0 (disagree), or NA (not applicable). The Actionability domain has 7 elements and they are all scored the same way. Total scores for both areas range from 0% to 100%, with higher percentages indicating information that is more understandable and useful for guiding patient behavior.[25]

Statistical Analysis

Descriptive statistics were presented as means, standard deviations, medians and interquartile ranges. Data distribution was evaluated using the Shapiro-Wilk test. Overall differences among AI models were assessed using the Kruskal-Wallis H test. When significant, pairwise post hoc comparisons were performed with Bonferroni correction because the assumptions of normality were not met. The Spearman's rank correlation coefficient was used to assess the associations between variables. Statistical analyses were performed using IBM SPSS Statistics version 27.

RESULTS

Kruskal-Wallis tests were used to compare differences in response characteristics by type of AI. There were significant differences between AI types for CLEAR, mGQS, accuracy, usefulness, FKGL and understandability% scores ($p<0.05$). Bonferroni post hoc tests showed that both Copilot and ChatGPT scored significantly higher on CLEAR than Gemini 3 ($p=0.001$ and $p<0.001$). For mGQS and Accuracy, Copilot was much higher than DeepSeek V3.1 and Gemini 3 ($p<0.001$). Copilot scored significantly lower and therefore more favorable Usefulness scores than DeepSeek V3.1 and Gemini 3 ($p<0.001$). Gemini 3 had a significantly higher FKGL score than DeepSeek V3.1 ($p=0.027$) and DeepSeek V3.1 was significantly better than Gemini 3 for understandability% ($p=0.040$) (**Table 1**).

Then, Spearman correlation analyses were performed separately for each AI model. In the case of ChatGPT-4, the CLEAR score was highly and positively correlated with mGQS and Accuracy ($r=0.856$, $p<0.05$) but highly and negatively correlated with Usefulness ($r=-0.856$, $p<0.05$). mGQS was found to be perfectly and positively correlated with accuracy

Table 1. Distribution and comparison of features of responses according to AI types

	ChatGPT-4		Copilot		DeepSeek V3.1		Gemini 3			
	Mean±SD	Median (25%-75% q)	Mean±SD	Median (25%-75% q)	Mean±SD	Median (25%-75% q)	Mean±SD	Median (25%-75% q)	Test statistics	p
CLEAR	24±0.86	24 (24-24.5)	24.2±0.89	24 (24-25)	23.25±1.77	24 (24-24)	23.05±0.51	23 (23-23)	24.299	<0.001*
mGQS	4.25±0.44	4 (4-4.5)	4.5±0.51	4.5 (4-5)	4±0	4 (4-4)	4±0	4 (4-4)	22.282	<0.001*
Accuracy	4.25±0.44	4 (4-4.5)	4.5±0.51	4.5 (4-5)	4±0	4 (4-4)	4±0	4 (4-4)	22.282	<0.001*
Usefulness	1.75±0.44	2 (1.5-2)	1.5±0.51	1.5 (1-2)	2±0	2 (2-2)	2±0	2 (2-2)	22.282	<0.001*
FRE	53.95±17.68	57.4 (44.55-66.4)	47.47±12.46	46.7 (37.75-59.1)	56.11±20.89	55.3 (39.75-62.85)	46.48±13.54	47.8 (37.25-56.35)	4.311	0.230
FKGL	8.48±2.8	8.35 (7.25-9.5)	9.81±2.24	10.25 (8.05-11.25)	7.32±3.43	7.4 (6.6-9.35)	10.82±2.23	10.7 (9.15-12.1)	16.868	<0.001*
DISCERN	27.7±0.98	28 (28-28)	28±2.15	28 (27.5-28)	27.5±1.54	28 (27.5-28)	27.35±1.87	28 (28-28)	0.404	0.939
U%	75.05±18.76	83 (67-87)	72.15±7.21	75 (67-79)	78.75±10.19	82.5 (76-85.5)	70.35±9.94	70 (66-77)	11.322	0.010*
A%	57±17.5	60 (40-70)	61±17.74	60 (60-70)	54.9±21.59	58.5 (40-73.5)	54±9.4	60 (40-60)	1.905	0.592

*p<0.05, AI: Artificial Intelligence, SD: Standard deviation, q: Quartile, mGQS: Modified Global Quality Score, FRE: Flesch reading ease, FKGL: Flesch-Kincaid grade level, U: Understandability, A: Actionability (r=1.000, p<0.05) and both were found to be perfectly and negatively correlated with usefulness (r=-1.000, p<0.05). FRE and FKGL were very strongly correlated (r=-0.967, p<0.05). There was a moderate negative correlation between FKGL and DISCERN (r=-0.456, p<0.05). Correlation between understandability% and actionability% was moderate (r=0.587, p<0.05) and DISCERN with understandability% and actionability% was moderate (r=0.585 and r=0.537, p<0.05) (**Table 2**).

Regarding Copilot, CLEAR score was strongly and positively correlated with mGQS and accuracy (r=0.877, p<0.05), and strongly and negatively correlated with usefulness (r=-0.877, p<0.05). It also indicated a moderate positive correlation with understandability% (r=0.485, p<0.05). mGQS and accuracy were perfectly correlated (r=1.000, p<0.05) and both were perfectly negatively correlated with usefulness (r=-1.000, p<0.05). FRE was very strongly and negatively correlated to FKGL (r=-0.908, p<0.05). FKGL was moderately negatively correlated to DISCERN (r=-0.579, p<0.05). DISCERN had moderate positive correlations with understandability% and actionability% (r=0.600 and r=0.544, p<0.05), and understandability% and actionability% were also moderately positively correlated (r=0.455, p<0.05) (**Table 3**).

In DeepSeek V3.1, CLEAR score showed a moderate positive correlation with the actionability% (r=0.589, p<0.05). mGQS had perfect positive correlations with accuracy and usefulness (r=1.000, p<0.05). Accuracy and usefulness were perfectly correlated (r=1.000). FRE and FKGL were highly negatively correlated (r=-0.971, p<0.05). There was a moderate positive correlation between understandability% and DISCERN (r=0.499, p<0.05) and understandability% and actionability% (r=0.640, p<0.05) (**Table 4**).

For Gemini 3, mGQS was perfectly positively correlated with both accuracy and usefulness (r=1.000, p<0.05) and accuracy was also perfectly positively correlated with usefulness (r=1.000,

Table 2. Relationships between metrics for ChatGPT-4

		mGQS	Accuracy	Usefulness	FRE	FKGL	DISCERN	U%	A%
CLEAR	r	0.856	0.856	-0.856	-0.161	0.248	0.094	0.248	0.360
	p	<0.001*	<0.001*	<0.001*	0.497	0.292	0.694	0.291	0.119
mGQS	r		1.000	-1.000	-0.250	0.351	-0.040	0.293	0.159
	p		<0.001*	<0.001*	0.287	0.130	0.868	0.210	0.504
Accuracy	r			-1.000	-0.250	0.351	-0.040	0.293	0.159
	p			<0.001*	0.287	0.130	0.868	0.210	0.504
Usefulness	r				0.250	-0.351	0.040	-0.293	-0.159
	p				0.287	0.130	0.868	0.210	0.504
FRE	r				1.000	-0.967	0.420	0.225	-0.138
	p					<0.001*	0.065	0.341	0.561
FKGL	r						-0.456	-0.218	0.099
	p						0.043*	0.357	0.677
DISCERN	r							0.585	0.537
	p							0.007*	0.015*
Understandability%	r								0.587
	p								0.007*

*p<0.05 and r: Correlation coefficient, mGQS: Modified Global Quality Score, FRE: Flesch reading ease, FKGL: Flesch-Kincaid grade level, U: Understandability, A: Actionability

Table 3. Relationships between metrics for Copilot

		mGQS	Accuracy	Usefulness	FRE	FKGL	DISCERN	U%	A%
CLEAR	r	0.877	0.877	-0.877	-0.309	0.051	0.322	0.485	0.268
	p	<0.001*	<0.001*	<0.001*	0.185	0.831	0.166	0.030*	0.253
mGQS	r		1.000	-1.000	-0.434	0.200	0.059	0.315	0.096
	p		<0.001*	<0.001*	0.056	0.399	0.805	0.176	0.689
Accuracy	r			-1.000	-0.434	0.200	0.059	0.315	0.096
	p			<0.001*	0.056	0.399	0.805	0.176	0.689
Usefulness	r				0.434	-0.200	-0.059	-0.315	-0.096
	p				0.056	0.399	0.805	0.176	0.689
FRE	r				1.000	-0.908	0.357	0.085	0.121
	p					<0.001*	0.122	0.722	0.612
FKGL	r						-0.579	-0.310	-0.345
	p						0.007*	0.183	0.137
DISCERN	r							0.600	0.544
	p							0.005*	0.013*
Understandability%	r								0.455
	p								0.044*

*p<0.05 and r: Correlation coefficient, mGQS: Modified Global Quality Score, FRE: Flesch reading ease, FKGL: Flesch-Kincaid grade level, U: Understandability, A: Actionability p<0.05). FRE was strongly and negatively correlated with FKGL (r=-0.964, p<0.05), and was moderately and positively correlated with understandability% (r=0.530, p<0.05). FKGL was moderately correlated with understandability% (r=-0.539, p<0.05). Finally, understandability% and actionability% were strongly positively correlated (r=0.778, p<0.05) (**Table 5**).

DISCUSSION

This study is among the first to systematically identify the clinically relevant patient questions about apical root-end resection and to evaluate the accuracy, content quality, readability, and patient-centered usability of responses generated by four AI-powered conversational agents (ChatGPT-4, DeepSeek V3.1, Copilot, and Gemini 3). Our findings revealed significant differences in response quality, accuracy, and readability across the evaluated platforms. Models differed overall in terms of readability. Some responses were suitable for patient education and some were beyond the recommended grade level. There were also limitations in the accuracy and depth of content. The results are consistent with the previous literature emphasizing the potential and limits of AI-based tools in dentistry. Compared to prior research that mainly assesses readability metrics,[16,18,23,26] the application of the PEMAT and DISCERN tools in our study provides a more complex assessment of whether content produced by AI is both readable and actionable and of high quality from a patient education standpoint.

Bahadir et al.[27] found that patients have a positive attitude towards AI-assisted dental diagnostics and trust their dentists. Similarly, Kosan et al.[28] showed that patients tend

Table 4. Relationships between metrics for DeepSeek V3.1

		mGQS	Accuracy	Usefulness	FRE	FKGL	DISCERN	U%	A%
CLEAR	r	0.427	0.427	0.427	0.228	-0.274	0.065	0.437	0.589
	p	0.061	0.061	0.061	0.334	0.242	0.785	0.054	0.006*
mGQS	r		1.000	1.000	0.139	-0.338	0.442	0.342	0.183
	p		<0.001*	<0.001*	0.558	0.145	0.051	0.139	0.440
Accuracy	r			1.000	0.139	-0.338	0.442	0.342	0.183
	p			<0.001*	0.558	0.145	0.051	0.139	0.440
Usefulness	r				0.139	-0.338	0.442	0.342	0.183
	p				0.558	0.145	0.051	0.139	0.440
FRE	r					-0.971	0.161	0.042	0.243
	p					<0.001*	0.499	0.861	0.301
FKGL	r						-0.242	-0.075	-0.183
	p						0.303	0.752	0.440
DISCERN	r							0.499	0.227
	p							0.025*	0.335
U%	r								0.640
	p								0.002*

*p<0.05 and r: Correlation coefficient, mGQS: Modified Global Quality Score, FRE: Flesch reading ease, FKGL: Flesch-Kincaid grade level, U: Understandability, A: Actionability

Table 5. Relationships between metrics for Gemini 3

		mGQS	Accuracy	Usefulness	FRE	FKGL	DISCERN	U%	A%
CLEAR	r	0.026	0.026	0.026	-0.211	0.298	-0.018	-0.404	-0.374
	p	0.913	0.913	0.913	0.371	0.202	0.939	0.077	0.104
mGQS	r		1.000	1.000	-0.060	-0.020	0.392	0.083	-0.150
	p		<0.001*	<0.001*	0.803	0.934	0.087	0.729	0.527
Accuracy	r			1.000	-0.060	-0.020	0.392	0.083	-0.150
	p			<0.001*	0.803	0.934	0.087	0.729	0.527
Usefulness	r				-0.060	-0.020	0.392	0.083	-0.150
	p				0.803	0.934	0.087	0.729	0.527
FRE	r					-0.964	0.225	0.530	0.435
	p					0.001*	0.339	0.016*	0.055
FKGL	r						-0.233	-0.539	-0.388
	p						0.322	0.014*	0.091
DISCERN	r							0.309	0.298
	p							0.186	0.201
U%	r								0.778
	p								<0.001*

*p<0.05 and r: Correlation coefficient, mGQS: Modified Global Quality Score, FRE: Flesch reading ease, FKGL: Flesch-Kincaid grade level, U: Understandability, A: Actionability to have a positive attitude towards AI in dentistry and that AI-assisted diagnostics could enhance patients' caries detection on radiographs without significantly affecting their trust in dental professionals. Previous studies[27,28] have reported generally positive patient attitudes towards AI in dentistry, although they stress the continued importance of trust in dental professionals. The findings underscore the importance of AI-generated information being accurate and understandable, and used as a supplement, not a substitute, to professional consultation. Thus, AI-powered conversational agents seem to have increasing potential in both clinical education and patient information. In this context, we sought to evaluate the accuracy of AI responses on apical root-end resection, a procedure where patients, even with detailed in-office explanations, frequently retain residual uncertainty.

Previous studies indicate that AI systems can achieve remarkable accuracy in areas such as dental implant prognosis, peri-implant bone loss detection, and apical surgery. For instance, Alqutaibi et al.[29] described that AI algorithms predicting implant prognosis from radiographic data achieved up to 99.8% accuracy, while Mugri et al.[30] reported that AI models detecting peri-implant bone loss demonstrated 61-94.7% accuracy with high specificity. While the high performance of AI in image-based dental diagnostic tasks has been demonstrated, such findings are not directly comparable with the evaluation of generative chatbot responses. Generative models are evaluated separately, as they could generate linguistically plausible but clinically incorrect information. In endodontics, Durmazpınar et al.[31] showed that the diagnostic accuracy of ChatGPT-4 was higher than that of 3rd- and 5th-year dental students, suggesting the possibility of AI as a clinical education tool. Ekmekci et al.[32] also reported that ChatGPT-4 with a PDF plugin achieved higher accuracy than ChatGPT-4o and Gemini in answering questions about regenerative endodontic treatment. The studies support the proposition of AI generating clinically relevant and accurate responses, consistent with our finding that all four chatbots scored high on the CLEAR, mGQS, and Accuracy scales. Our query set was created from established clinical consensus statements and the AAE and Cohen's Pathways of the Pulp guidelines, thus establishing an evidence-based reference standard based on established clinical guidelines and consensus sources in evaluating the AI models, rather than arbitrary patient inquiries.[4,5,10] We find more subtle differences across platforms as well. Esmailpour et al.[14] found that Gemini is better in factual correctness and ChatGPT-4 and Copilot showed more uniform overall performance in a broader set of questions. There was no difference in DISCERN scores between the four models. However, Copilot scored the highest for mGQS and Accuracy. These findings partly align with previous work[33] that indicates AI-generated answers may offer high-quality patient information overall but are still not deep enough for clinical decision-making. Copilot scored the highest mean CLEAR score followed by ChatGPT-4, both significantly higher than Gemini 3, suggesting that readability and understandability may not fully align with factual accuracy, which is also noted in the evaluation of AI responses to emergency care questions by Yau et al.[34] Gemini 3 had a higher FKGL and lower understandability than DeepSeek V3.1, suggesting that increased linguistic complexity may limit the accessibility of otherwise informative responses. The information in Gemini 3 was generally accurate, but the higher reading grade level appeared to compromise practical understandability. Importantly, the negative correlations between Usefulness and CLEAR, mGQS, or Accuracy should not be interpreted as indicating that more accurate or higher-quality responses were less useful. This pattern primarily resulted from the opposite scoring directions: lower Usefulness scores denoted greater usefulness, whereas higher CLEAR, mGQS, and Accuracy scores indicated better performance. Because the Usefulness scale was reverse-coded, with lower scores indicating greater usefulness, negative correlations with positively oriented measures such as CLEAR, mGQS, and Accuracy represented concordant rather than contradictory performance.

A multidisciplinary panel, which included periodontists and an oral and maxillofacial radiologist, broadly assessed the AI-generated responses. This multidisciplinary approach enabled evaluation of the responses from the perspectives of the periodontist/surgeon and the radiologist relevant to apical root-end resection and follow-up.

In our study, the analysis of readability indicated that the responses from Gemini 3 and Copilot had higher FKGL scores, showing that generally, a higher educational reading level was required for responses. DeepSeek V3.1 had the lowest mean FKGL score. Büker and Mercan[26] reported that none of the AI chatbots met the recommended level of readability for patient education materials, but Gemini was superior to the other models in terms of readability, accuracy and quality of response. Dursun and Geçer[35] found that all AI chatbots provided generally accurate and moderate-to-good quality information on clear aligners, but the responses from Gemini were significantly more readable than ChatGPT-3.5, ChatGPT-4 and Copilot. Özcivelek and Özcan[17] found that the readability of AI chatbots was different when patients asked questions about dental and maxillofacial prostheses. The best readability was DeepSeek-R1 and the worst readability was Dental GPT. Esmailpour et al.[14] evaluated the performance of AI chatbots in responding to patients' questions regarding dental prosthesis. They observed that Google Gemini scored the highest quality scores, whereas ChatGPT offered more complex and less readable responses. Variability in readability across studies may be attributed to differences in model architecture, training data, prompt design, domain-specific fine-tuning, and question complexity. The choice of the AI architecture, the training datasets, the prompt design and the domain-specific fine-tuning can impact the complexity of the sentence, the choice of words and the overall structure of the text. Also, the type of question asked, ranging from a technical clinical procedure to a general patient concern, will probably affect the language complexity required to give an accurate, yet comprehensible, answer. Together, these factors help to explain why some AI outputs have greater readability and others are above the recommended levels for patient education materials.

Limitations

This study has a few limitations. First, only four AI-powered chatbots were assessed (ChatGPT-4, DeepSeek V3.1, Copilot, and Gemini 3), which limits the generalizability of the findings to other AI models or future versions. Second, the responses were collected on one day and only the initial responses were considered. As AI models are constantly updating and improving their responses, the results might be different if the same questions were asked at a different time or with different prompts. Third, although we systematically assessed the readability by established indices, the patient-centered understandability and actionability were assessed by the PEMAT tool and not by testing on patients, which may not fully reflect the real-world comprehension. Finally, the complexity and wording of questions could have affected the quality and clarity of the AI-generated responses. Together, these limitations highlight the need for careful interpretation and the need for further studies to assess AI chatbots in other languages, clinical domains, patient populations, and longitudinal settings.

CONCLUSION

The tested models presented considerable potential for the delivery of generally accurate and clear information regarding apical root-end resection, demonstrating different performance profiles. Copilot had the highest overall mGQS and Accuracy scores, the most favorable usefulness scores, and the highest mean CLEAR score, with ChatGPT-4 close behind, both significantly outperforming Gemini 3 on CLEAR. There was no significant difference in the DISCERN scores between the four models, indicating comparable performance for this measure. Gemini 3 scored higher on FKGL and lower on understandability compared with DeepSeek V3.1, suggesting that clinically informative content may still be difficult to understand for patients when presented at a higher reading level. Because the usefulness scale was reverse-coded, negative correlations with positively oriented measures such as CLEAR, mGQS and Accuracy reflected concordant rather than contradictory performance. Overall, the findings suggest that technical accuracy, readability, understandability and actionability are related but distinct dimensions of patient information generated by AI. AI chatbots can be useful additions to answering patient questions, but clinician oversight is advised to make sure that the information provided is clinically accurate, understandable, and practically actionable. Future developments should therefore take into account the trade-off between technical accuracy and patient-centered communication as well as existing health literacy standards.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study did not require ethical approval as it did not involve any human subjects or animal experiments.

Informed Consent

Because the study did not involve human participants, written informed consent was not required.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: TP, SG, SCŞ, ÖSA, GU, ŞÇB, SÖ; Design: TP, SG, SCŞ, ÖSA, GU, ŞÇB, SÖ; Control: TP, SG, SCŞ, ÖSA, GU, ŞÇB, SÖ; Resources: TP, SG; Materials: TP, SG; Data Collection and/or Processing: TP, SG, SCŞ, ÖSA, GU, ŞÇB, SÖ; Analysis and/or Interpretation: TP, SG, SCŞ, ÖSA, GU; Literature Review: SCŞ, ÖSA, GU, ŞÇB, SÖ; Writing the Article: TP, SG, SCŞ, ÖSA, GU, ŞÇB, SÖ; Critical Review: TP, SCŞ, ÖSA, GU.

REFERENCES

1. Möller AJ, Fabricius L, Dahlén G, et al. Influence on periapical tissues of indigenous oral bacteria and necrotic pulp tissue in monkeys. *Scand J Dent Res.* 1981;89(6):475-484. doi:10.1111/j.1600-0722.1981.tb01711.x
2. Song M, Nam T, Shin SJ, Kim E. Comparison of clinical outcomes of endodontic microsurgery: 1 year versus long-term follow-up. *J Endod.* 2014;40(4):490-494. doi:10.1016/j.joen.2013.10.034
3. Setzer FC, Shah SB, Kohli MR, Karabucak B, Kim S. Outcome of endodontic surgery: a meta-analysis of the literature-part 1: comparison of traditional root-end surgery and endodontic microsurgery. *J Endod.* 2010;36(11):1757-1765. doi:10.1016/j.joen.2010.08.007
4. American Association of Endodontists. Glossary of Endodontic Terms. 10th ed. American Association of Endodontists; 2020.
5. Berman LH, Hargreaves KM. Cohen's Pathways of the Pulp. 12th ed. Elsevier; 2020.
6. Jiang F, Jiang Y, Zhi H, et al. Artificial Intelligence in healthcare: past, present and future. *Stroke Vasc Neurol.* 2017;2(4):230-243. doi:10.1136/svn-2017-000101
7. Schwendicke F, Samek W, Krois J. Artificial Intelligence in dentistry: chances and challenges. *J Dent Res.* 2020;99(7):769-774. doi:10.1177/0022034520915714
8. Keleş A, Keskin C. A micro-computed tomographic study of band-shaped root canal isthmuses, having their floor in the apical third of mesial roots of mandibular first molars. *Int Endod J.* 2018;51(2):240-246. doi:10.1111/iej.12842
9. Xu T, Fan W, Tay FR, et al. Micro-computed tomographic evaluation of the prevalence, distribution, and morphologic features of accessory canals in Chinese permanent teeth. *J Endod.* 2019;45(8):994-999. doi:10.1016/j.joen.2019.04.001
10. Ordinola-Zapata R, Martins JNR, Niemczyk S, Bramante CM. Apical root canal anatomy in the mesiobuccal root of maxillary first molars: influence of root apical shape and prevalence of apical foramina-a micro-CT study. *Int Endod J.* 2019;52(8):1218-1227. doi:10.1111/iej.13109
11. Xu T, Tay FR, Gutmann JL, et al. Micro-computed tomography assessment of apical accessory canal morphologies. *J Endod.* 2016;42(5):798-802. doi:10.1016/j.joen.2016.02.006
12. Stropko JJ, Doyon GE, Gutmann JL. Root-end management: resection, cavity preparation, and material placement. *Endod Topics.* 2005;11(1):131-151. doi:10.1111/j.1601-1546.2005.00158.x
13. 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. doi:10.3390/healthcare10071269
14. Esmailpour H, Rasaie V, Babaee Hemmati Y, Falahchai M. Performance of Artificial Intelligence chatbots in responding to the frequently asked questions of patients regarding dental prostheses. *BMC Oral Health.* 2025;25(1):574. doi:10.1186/s12903-025-05965-9
15. 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(10):1239-1245. doi:10.1016/j.joms.2024.06.177
16. Ceylan Şen S, Çardakci Bahar Ş, Saraç Atagün Ö, et al. Quality and reliability of AI information on dental implant failure: a comparative multi-model analysis. *J Craniofac Surg.* 2026;37(3-4):675-680. doi:10.1097/SCS.0000000000012415
17. Özcivelek T, Özcan B. Comparative evaluation of responses from DeepSeek-R1, ChatGPT-o1, ChatGPT-4, and dental GPT chatbots to patient inquiries about dental and maxillofacial prostheses. *BMC Oral Health.* 2025;25(1):871. doi:10.1186/s12903-025-06267-w
18. Ceylan Şen S, Saraç Atagün Ö, Ustaoğlu G, et al. Do AI chatbots tell the truth about dentin hypersensitivity? A comparative evaluation of quality, accuracy, and readability. *Cumhuriyet Dent J.* 2026;29(1):138-147. doi:10.7126/cumudj.1848545
19. Helvacioglu-Yigit D, Demirturk H, Ali K, et al. Evaluating Artificial Intelligence chatbots for patient education in oral and maxillofacial radiology. *Oral Surg Oral Med Oral Pathol Oral Radiol.* 2025;139(6):750-759. doi:10.1016/j.oooo.2025.01.001

20. National Library of Medicine. How to write easy-to-read health materials. Updated November 2020. Accessed December 12, 2022. https://medlineplus.gov/all_easytoread.html
21. Sallam M, Barakat M, Sallam M. Pilot testing of a tool to standardize the assessment of the quality of health information generated by Artificial Intelligence-based models. *Cureus.* 2023;15(11):e49373. doi:10.7759/cureus.49373
22. Bernard A, Langille M, Hughes S, et al. A systematic review of patient inflammatory bowel disease information resources on the World Wide Web. *Am J Gastroenterol.* 2007;102(9):2070-2077. doi:10.1111/j.1572-0241.2007.01325.x
23. Hatia A, Doldo T, Parrini S, et al. Accuracy and completeness of ChatGPT-generated information on interceptive orthodontics: a multicenter collaborative study. *J Clin Med*. 2024;13(3):735. doi:10.3390/jcm13030735
24. Yeo YH, Samaan JS, Ng WH, et al. Assessing the performance of ChatGPT in answering questions regarding cirrhosis and hepatocellular carcinoma. *Clin Mol Hepatol.* 2023;29(3):721-732. doi:10.3350/cmh.2023.0089
25. Alnsour MM, Alenezi R, Barakat M, et al. Assessing ChatGPT's suitability in responding to the public's inquires on the effects of smoking on oral health. *BMC Oral Health*. 2025;25(1):1207. doi:10.1186/s12903-025-06377-5
26. Büker M, Mercan G. Readability, accuracy and appropriateness and quality of AI chatbot responses as a patient information source on root canal retreatment: a comparative assessment. *Int J Med Inform*. 2025;201:105948. doi:10.1016/j.ijmedinf.2025.105948
27. Bahadir HS, Keskin NB, Çakmak EŞK, et al. Patients' attitudes toward Artificial Intelligence in dentistry and their trust in dentists. *Oral Radiol*. 2025;41(1):52-59. doi:10.1007/s11282-024-00775-1
28. Kosan E, Krois J, Wingenfeld K, et al. Patients' perspectives on Artificial Intelligence in dentistry: a controlled study. *J Clin Med*. 2022;11(8):2143. doi:10.3390/jcm11082143
29. Alqutaibi AY, Algabri RS, Alamri AS, et al. Advancements of Artificial Intelligence algorithms in predicting dental implant prognosis from radiographic images: a systematic review. *J Prosthet Dent*. 2025;134(6):2177-2188. doi:10.1016/j.prosdent.2024.10.036
30. Mugri MH. Accuracy of Artificial Intelligence models in detecting peri-implant bone loss: a systematic review. *Diagnostics (Basel)*. 2025;15(6):655. doi:10.3390/diagnostics15060655
31. Durmazpinar PM, Ekmekci E. Comparing diagnostic skills in endodontic cases: dental students versus ChatGPT-4o. *BMC Oral Health*. 2025;25(1):457. doi:10.1186/s12903-025-05857-y
32. Ekmekci E, Durmazpinar PM. Evaluation of different Artificial Intelligence applications in responding to regenerative endodontic procedures. *BMC Oral Health*. 2025;25(1):53. doi:10.1186/s12903-025-05424-5
33. Terzi M, Yavuz MC, Bicer T, Buyuk SK. Evaluation of Artificial Intelligence robot's knowledge and reliability on dental implants and peri-implant phenotype. *Sci Rep*. 2025;15(1):9519. doi:10.1038/s41598-025-94576-z
34. Yau JYS, Saadat S, Hsu E, et al. Accuracy of prospective assessments of 4 large language model chatbot responses to patient questions about emergency care: experimental comparative study. *J Med Internet Res*. 2024;26:e60291. doi:10.2196/60291
35. Dursun D, Bilici Geçer R. Can Artificial Intelligence models serve as patient information consultants in orthodontics? *BMC Med Inform Decis Mak*. 2024;24(1):211. doi:10.1186/s12911-024-02619-8

Journal of Dental Sciences and Education

Original Article

DOI: 10.51271/JDSE-0082

Comparative evaluation of AI-based chatbots in answering prosthodontic questions from the Turkish Dental Specialty Exam (DUS)[a]

Bülent Kadir Tartuk

Department of Prosthodontics, Faculty of Dentistry, Bingöl University, Bingöl, Turkiye

Cite this article as: Tartuk BK. Comparative evaluation of AI-based chatbots in answering prosthodontic questions from the Turkish Dental Specialty Exam (DUS). *J Dent Sci Educ.* 2026;4(3):84-90. doi:10.51271/JDSE-0082

Received: 01.06.2026 ◆ **Accepted**: 30.07.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Aims: This study compared the performance of the ChatGPT, Gemini, Copilot, and DeepSeek large language model-based chatbots in answering prosthodontic questions from the Turkish Dental Specialty Exam (DUS).

Methods: A total of 126 prosthodontic questions from DUS examinations conducted between 2012 and 2021 were included in the study after excluding visually based and canceled questions. The questions were classified as knowledge-based and case-based. Additionally, they were categorized according to topic areas as fixed prosthodontics, removable prosthodontics and implant-supported prosthodontics. The data were analyzed using the generalized estimating equations model, which accounts for repeated measurements, and differences between models were evaluated using Bonferroni-corrected post hoc tests.

Results: All chatbot models included in this study achieved an overall accuracy rate above 80% (ChatGPT-4o: 85.71%; Copilot: 83.33%; Gemini 3 Flash: 81.75%; DeepSeek-V3.2: 80.16%). No statistically significant difference was observed among the models in terms of overall performance (p=0.619). All models demonstrated significantly higher performance on knowledge-based questions than on case-based questions (p<0.05). Topic-based analysis showed numerically lower accuracy for implant-supported prosthodontic questions than for fixed and removable prosthodontic questions. However, this finding should be interpreted cautiously because the implant-supported subgroup included only four questions. No significant relationship was found between examination years and the accuracy rates of the models (χ^2=7.1; p=0.62).

Conclusion: The examined Artificial Intelligence models demonstrated a high level of accuracy in answering prosthodontic specialty questions. These systems have the potential to serve as supportive tools in dental education and examination preparation processes. However, owing to their limitations in questions requiring clinical reasoning, particularly case-based questions, their use should be approached cautiously and critically.

Keywords: Artificial Intelligence, large language models, prosthodontics, DUS, dental education

[a]This study was presented orally at the '4th National İstanbul Aydın University Faculty of Dentistry Young Dentists and Student Congress'.

INTRODUCTION

Artificial Intelligence (AI) refers to a field of technology aimed at developing computer systems and algorithms capable of performing tasks that typically require human intelligence.[1] AI applications based on methods such as deep learning and neural networks have led to substantial transformations in numerous fields, including healthcare.[2] In recent years, these technologies have enabled major advances in healthcare services by facilitating rapid data analysis, early error detection, and support for clinical decision-making processes, thereby contributing to improved workflow efficiency, increased productivity, and reduced error rates in the healthcare sector.[3,4]

The rapid advancement of AI in the fields of health sciences and education has also significantly transformed access to information and academic learning methods.[5] One of the most notable outcomes of this transformation has been the emergence of large language models (LLMs), advanced AI systems capable of mimicking human language comprehension, interpretation, and generation.[6] These models are trained on large and heterogeneous text corpora using deep learning algorithms and are capable of producing highly human-like, contextually relevant outputs.[7,8]

LLM-based conversational systems such as ChatGPT, Gemini, Copilot, and DeepSeek, which represent some of the most widely accessible applications of this technology today, are capable of successfully performing complex tasks including text generation, summarization, translation, and question answering.[9] Nevertheless, LLMs may differ in terms of architecture, training data sources, updating mechanisms, and multimodal capabilities, which can lead to variations in accuracy and response quality across disciplines and question types.[10] Therefore, particularly in fields requiring specialized expertise such as medicine and dentistry, the accuracy and contextual reliability of the content generated by these models should be evaluated carefully and systematically.[11,12]

The integration of AI-powered chatbots into dental education and clinical practice may only be feasible to the extent

Corresponding Author: Bülent Kadir Tartuk, bktartuk@bingol.edu.tr

that these systems can generate sufficiently accurate and reliable responses within relevant areas of expertise.[2] In this context, standardized examinations requiring high levels of cognitive competence provide an important benchmark for evaluating the performance of AI systems. In Turkiye, the Dental Specialty Exam (DUS) is a nationally administered, high-stakes examination used as the primary criterion for admission to postgraduate dental specialty training programs. The examination consists of multiple-choice questions designed to assess dentists' professional knowledge in both clinical dentistry and basic sciences.[2,13] Evaluating the performance of AI-based chatbots on this examination may provide valuable insight into the potential applications of these systems in dentistry.[14]

In recent years, the number of studies evaluating the performance of AI-based LLMs through the DUS and similar examinations has increased considerably. However, a substantial proportion of these studies have focused on specific specialty fields, limited question sets, or selected model comparisons. Therefore, further comparative evaluations of different models within specific dental disciplines are needed.

Prosthodontics is a complex dental specialty that encompasses biomechanical principles, materials science, occlusion, esthetic planning, and clinical decision-making processes.[15] Owing to its multidimensional nature, prosthodontics provides an appropriate testing domain for assessing the ability of LLMs to integrate theoretical knowledge with clinical reasoning. In particular, the limited number of studies investigating the comparative performance of different AI models in prosthodontics according to examination years, question types, and subtopic categories highlights the need for more comprehensive evaluations in this field.

Accordingly, this study aimed to evaluate and compare the performance of AI-powered chatbots, including ChatGPT, Gemini, Copilot, and DeepSeek, in answering prosthodontic questions from previously administered DUS examinations. The null hypotheses tested were as follows: no significant difference among the models in terms of overall accuracy rates (H_{01}), no performance difference between knowledge-based and case-based questions (H_{02}), and similar accuracy rates across removable, fixed, and implant-supported prosthodontic questions (H_{03}).

METHODS

Ethics

Ethical approval was not required for this study because all data were obtained from publicly available examination questions published on the official website of the Measurement, Selection and Placement Center (ÖSYM). Informed consent was also not required, because the study did not involve human participants, personal data, or clinical records. All procedures were carried out in accordance with the ethical rules and the principles.

Data Source and Eligibility Criteria

The DUS is a nationwide centralized examination consisting of 120 multiple-choice questions, including 40 basic sciences and 80 clinical sciences questions, 10 of which are related to prosthodontics. First administered in 2012, the DUS was conducted twice annually until 2014 and once annually between 2015 and 2021. Accordingly, a total of 13 examinations containing prosthodontic questions were included in the present study. Examinations conducted after ÖSYM discontinued the public release of DUS questions in 2021 were excluded.

The exclusion criteria comprised questions officially canceled by ÖSYM because of issues such as the presence of multiple correct answers or the absence of a valid correct answer, as well as questions containing figures, images, or graphical content. Questions with visual content were excluded because current AI models are predominantly based on text-processing architectures and still exhibit limitations in effectively interpreting complex visual information.[16]

Initially, 130 prosthodontic questions were identified during the study period. After excluding questions that met the exclusion criteria, 126 questions were included in the analysis.

Question Classification

To ensure analytical consistency, the included questions were independently evaluated by two prosthodontists and classified according to three criteria; the examination year, question type, and topic area. Based on the question type, the questions were categorized as either knowledge-based or case-based. Knowledge-based questions were defined as those requiring direct factual knowledge related to basic concepts, material properties, or theoretical information, whereas case-based questions were defined as those involving clinical scenarios requiring diagnosis, treatment planning, or clinical decision-making.

Of the 126 questions included in the overall analysis, 97 were assigned to one of the three predefined topic categories: fixed prosthodontics (n=48), removable prosthodontics (n=45), and implant-supported prosthodontics (n=4) (**Table 1**). The remaining 29 questions were excluded from the topic-based subgroup analysis because they addressed cross-cutting subjects, including impression materials, occlusion, and temporomandibular joint-related topics, that could not be assigned exclusively to a single prosthodontic treatment category. These questions were retained in the overall accuracy and question-type analyses.

Table 1. Classification of DUS prosthodontic questions according to topic area

Topic	Definition/scope
Fixed prosthesis	Questions related to diagnosis, treatment planning, tooth preparation principles, material selection, and occlusal considerations for fixed prosthetic restorations
Removable prosthesis	Questions concerning the planning, design, biomechanics, and clinical application of removable prosthetic treatments
Implant-supported prosthesis	Questions related to prosthetic rehabilitation using dental implants, including treatment planning, prosthetic design, and implant-supported restorations

DUS: Turkish Dental Specialty Exam

Query Protocol and Response Evaluation

All chatbot interactions were conducted on January 13, 2026, in a desktop environment using the Google Chrome browser (version 124.0; Google LLC, Mountain View, CA, USA) on the Windows 10 operating system (Microsoft Corp., Redmond, WA, USA). The final question set was presented in Turkish under standardized conditions using the publicly accessible versions of ChatGPT-4o (OpenAI, San Francisco, CA, USA; https://chat.openai.com/), Gemini 3 Flash (Google LLC, Mountain View, CA, USA; https://gemini.google.com/), Copilot (Microsoft Corp., Redmond, WA, USA; https://copilot.microsoft.com/), and DeepSeek-V3.2 (DeepSeek AI, Hangzhou, China; https://chat.deepseek.com/) available on that date. All systems were accessed through their official browser-based web platforms rather than application programming interfaces (APIs).

New accounts were created for each of the chatbots. In addition, a separate new session was initiated for each question to prevent the previous conversational context from influencing the AI system's responses. Before each question, the following standardized Turkish prompt was provided to all models: "assume that you are a candidate taking the DUS examination. Select the most appropriate answer from the options provided in the following multiple-choice questions. Each question contains five options, and only one option is correct."

Only the first response generated by each chatbot was included in the evaluation, and no regeneration of responses was performed. Responses were considered "correct" if they matched the official answer key published by the ÖSYM. The accuracy rate for each chatbot was calculated as the percentage obtained by dividing the number of correctly answered questions by the total number of questions answered.

Inter-rater reliability for the classification of question type, topic area, and response accuracy was assessed using Cohen's kappa (κ) coefficient. The Cohen's kappa value was 0.91, indicating a high level of agreement between evaluators. In cases of disagreement, a consensus was reached between the evaluators.

Statistical Analysis

All data analyses were performed using SPSS statistical software version 27.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were used to summarize the accuracy rates of the AI models. The dependent variable was defined as response accuracy (correct/incorrect). Because the responses were binary and each question was answered by all models, the generalized estimating equations (GEE) approach was employed to account for the repeated-measures structure of the data. A binomial distribution with a logit link function was used in the analyses, and the correlation structure was specified as exchangeable. The effects of model type, examination year, question type, and topic area on the probability of a correct response were evaluated. Bonferroni-corrected pairwise comparisons were performed for statistically significant findings. A p-value of <0.05 was considered statistically significant.

RESULTS

The distributions of correct and incorrect responses of the chatbot models according to question type are presented in **Table 2**. Across all models, the accuracy rates for knowledge-based questions were higher than those for case-based questions. Among the knowledge-based questions, ChatGPT demonstrated the highest accuracy rate (90.48%), followed by Copilot (87.62%), Gemini (86.67%), and DeepSeek (85.71%). For case-based questions, ChatGPT and Copilot achieved identical accuracy rates (61.90%). In contrast, Gemini (57.14%) and DeepSeek (52.38%) demonstrated lower accuracy rates for case-based questions.

Table 2. Accuracy rates of chatbot models according to question type

Model	Knowledge (n=105)		Case-based (n=21)		
	True	False	True	False	p-value
ChatGPT	90.48% (95)	9.52% (10)	61.90% (13)	38.10% (8)	0.031*
Gemini	86.67% (91)	13.33% (14)	57.14% (12)	42.86% (9)	0.039*
Copilot	87.62% (92)	12.38% (13)	61.90% (13)	38.10% (8)	0.025*
DeepSeek	85.71% (90)	14.29% (15)	52.38% (11)	47.62% (10)	0.028*

*Statistically significant

When all questions were evaluated together, ChatGPT achieved the highest overall accuracy rate (85.71%), followed by Copilot (83.33%), Gemini (81.75%) and DeepSeek (80.16%). However, no statistically significant differences were observed among the models in terms of knowledge-based questions, case-based questions, or overall accuracy rates (p=0.684, p=0.892, and p=0.619, respectively) (**Table 3**).

Table 3. Comparison of chatbot models according to question type and overall accuracy

		ChatGPT		Gemini		Copilot		DeepSeek		
Question type	n	True	False	True	False	True	False	True	False	p-value
Knowledge	105	90.48% (95)	9.52% (10)	86.67% (91)	13.33% (14)	87.62% (92)	12.38% (13)	85.71% (90)	14.29% (15)	0.684
Case-based	21	61.90% (13)	38.10% (8)	57.14% (12)	42.86% (9)	61.90% (13)	38.10% (8)	52.38% (11)	47.62% (10)	0.892
Total	126	85.71% (108)	14.29% (18)	81.75% (103)	18.25% (23)	83.33% (105)	16.67% (21)	80.16% (101)	19.84% (25)	0.619

In the Bonferroni-corrected pairwise comparison analyses, no statistically significant differences were detected among the chatbot models for either knowledge-based or case-based questions (p>0.05). This finding indicates that the models demonstrated comparable performance levels regardless of the question type (**Table 4**).

Table 4. Pairwise comparisons of chatbot performance according to question type

Chatbots	Knowledge (p)	Case-based (p)
ChatGPT-Gemini	0.342	0.743
ChatGPT-Copilot	0.514	1.000
ChatGPT-DeepSeek	0.287	0.612
Gemini-Copilot	0.621	0.721
Gemini-DeepSeek	0.418	0.534
Copilot-DeepSeek	0.552	0.674

In the topic-based evaluation, all the chatbot models showed similar accuracy rates for fixed and removable prosthodontic questions, whereas lower accuracy rates were observed for implant-supported prosthodontic questions. Within-model comparisons indicated a statistically significant effect of topic area on accuracy rates ($p<0.05$). However, the implant-supported prosthodontics subgroup comprised only four questions; therefore, the corresponding percentage estimates and subgroup comparisons should be regarded as unstable and interpreted cautiously (**Table 5**).

Table 5. Accuracy rates of chatbot models according to prosthodontic topic categories

	Fixed prosthesis (n=48)		Removable prosthesis (n=45)		Implant-supported prosthesis (n=4)		
Model	**True**	**False**	**True**	**False**	**True**	**False**	**p-value**
ChatGPT	87.50% (42)	12.50% (6)	86.67% (39)	13.33% (6)	50.00% (2)	50.00% (2)	0.021*
Gemini	85.42% (41)	14.58% (7)	84.44% (38)	15.56% (7)	50.00% (2)	50.00% (2)	0.027*
Copilot	83.33% (40)	16.67% (8)	82.22% (37)	17.78% (8)	25.00% (1)	75.00% (3)	0.018*
DeepSeek	81.25% (39)	18.75% (9)	80.00% (36)	20.00% (9)	25.00% (1)	75.00% (3)	0.015*

*Statistically significant

However, Bonferroni-corrected pairwise comparison analyses revealed no statistically significant differences among the chatbot models in any of the fixed, removable, or implant-supported prosthodontic categories ($p>0.05$) (**Table 6**).

Table 6. Pairwise comparisons of chatbot performance according to prosthodontic topic categories

Chatbots	Fixed prosthesis (p)	Removable prosthesis (p)	Implant-supported prosthesis (p)
ChatGPT-Gemini	0.79	0.82	1.00
ChatGPT-Copilot	0.56	0.59	0.50
ChatGPT-DeepSeek	0.34	0.39	0.50
Gemini-Copilot	0.71	0.74	0.50
Gemini-DeepSeek	0.49	0.54	0.50
Copilot-DeepSeek	0.76	0.79	1.00

The accuracy rates of the chatbot models according to the examination years are presented in **Figure**. Although some fluctuations in accuracy rates were observed across the years, statistical analysis demonstrated that the accuracy rates did not differ significantly according to the examination year ($\chi^2=7.1$; $p=0.62$). This finding suggests that chatbot performance remained consistent throughout the study period.

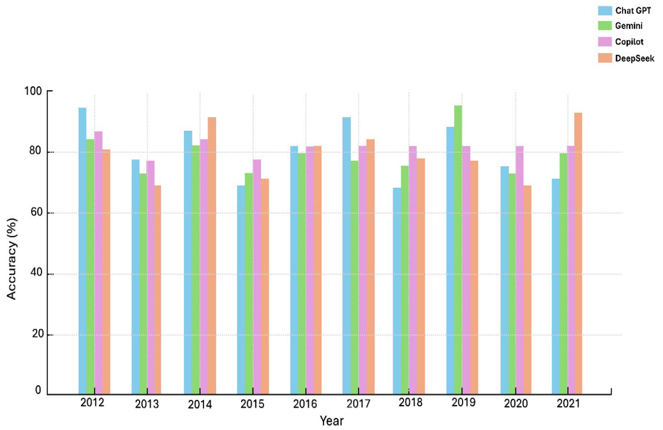

Figure. Accuracy rates of chatbot models across DUS examination years (2012-2021)

DUS: Turkish Dental Specialty Exam

DISCUSSION

In the present study, the accuracy of the responses generated by different AI-based LLMs to prosthodontic questions from the DUS was evaluated and compared. The findings demonstrated no statistically significant differences among the models in terms of overall accuracy. In addition, higher accuracy rates were observed for knowledge-based questions than for case-based questions. Topic-based evaluation revealed similar performance in fixed and removable prosthodontics, whereas numerically lower accuracy was observed in the implant-supported subgroup. Accordingly, the null hypothesis regarding overall performance differences among the models (H_{01}) was not rejected, whereas the null hypotheses concerning differences between knowledge-based and case-based questions (H_{02}) and differences according to topic area (H_{03}) were rejected.

The impact of AI has become increasingly prominent in healthcare fields such as medicine and dentistry, with applications in numerous processes including diagnosis, treatment planning, and patient follow-up.[17,18] Parallel to the growing adoption of AI-based applications in healthcare services, new opportunities have emerged in education to support teaching and learning processes for both students and educators.[19] In a previous study evaluating the performance of AI-based LLMs in medical education, the system was reported to achieve near-passing performance on the United States Medical Licensing Examination (USMLE) without specialized medical training or additional reinforcement.[20]

All models achieved overall accuracy rates above 80%, indicating potential value for examination preparation and information access. Nevertheless, these systems were not specifically developed for dental education or clinical decision-making, and high-test accuracy should not be interpreted as equivalent to domain-specific expertise.[2]

Consistent with our findings, previous studies have shown that LLMs such as ChatGPT can achieve passing-level or high accuracy rates on medical licensing examinations.[21] Recent investigations evaluating examination performance in various dental disciplines have further demonstrated that AI models can, in many cases, produce results comparable to human performance and occasionally even outperform human participants. These findings indicate that AI models may be effectively utilized to support educational materials and examination preparation processes.[2,22]

However, all chatbot models demonstrated higher accuracy for knowledge-based questions than for case-based questions, indicating that the models performed better on theoretical content than on clinically oriented reasoning. This finding is consistent with previous studies reporting that, although LLMs can achieve high levels of factual accuracy, they may have limitations in case interpretation and practice-based decision-making processes.[22] No statistically significant differences were detected across examination years, suggesting generally stable performance across the evaluated periods.

Ekici et al.[23] reported in their evaluation of DUS questions that AI-based LLMs performed better on knowledge-based questions, such as those in basic sciences, whereas their performance was lower on questions requiring not only factual knowledge but also clinical reasoning, discussion, and interpretation, such as those in the clinical sciences.

Similarly, Haberal et al.[24] evaluated the performance of different AI chatbots using restorative dentistry questions from the DUS and reported that all models achieved high accuracy levels. However, no statistically significant differences were found among the models in terms of overall accuracy rates. The study also noted a decrease in accuracy, particularly for complex questions that require clinical interpretation. This finding supports the view that LLMs may have limitations in situations requiring clinical decision-making and advanced reasoning.

The lower accuracy observed in case-based questions suggests that LLMs may have limitations in multistep decision-making processes that require clinical reasoning skills. Case-based questions require not only the recall of factual knowledge but also the integrated interpretation of clinical findings, formulation of differential diagnoses, and development of context-appropriate treatment plans. Previous studies have reported that LLMs often generate responses through probabilistic pattern matching based on training data rather than genuine clinical reasoning, and may demonstrate limited flexible reasoning, particularly in complex clinical scenarios. Moreover, these models have been reported to produce overconfident yet incorrect responses and to exhibit a tendency toward hallucination during clinical problem-solving. This may explain the reduced performance observed in the case-based questions.[25]

Numerically lower accuracy was observed for implant-supported prosthodontic questions. Implant dentistry involves multidisciplinary considerations, including prosthetic design, biomechanical assessment, anatomical risk analysis, and long-term complication management.[26-28] However, this finding should be interpreted cautiously because only four implant-supported prosthodontic questions were included in the present study. With such a small denominator, each response changed the subgroup accuracy estimate by 25 percentage points. Therefore, the observed difference may reflect sampling variability rather than a robust topic-specific performance deficit. Larger and more balanced question sets are required before firm conclusions can be drawn regarding the performance of LLMs in implant-supported prosthodontics.

Turan Gökduman et al.[29] evaluated DUS endodontic questions and reported accuracy rates of 82.8%, 83.6%, and 77.9% for ChatGPT, Gemini, and Copilot, respectively. No statistically significant overall difference was found among the models. In contrast to the present study, Copilot demonstrated a significantly higher accuracy rate during 2012–2015 than during the 2016–2021 period. These findings suggest that the performance of AI models across different dental disciplines within the DUS may be broadly similar in terms of overall accuracy, although performance fluctuations may occur in specific models depending on the examination year and question structure.

A study evaluating oral radiology questions from the DUS in Turkiye compared the performance of ChatGPT and Copilot and reported that ChatGPT achieved a markedly higher accuracy than Copilot. In that study, ChatGPT reached an accuracy rate of 86.1%, whereas Copilot achieved 41.5%. This finding suggests that ChatGPT may outperform Copilot in answering multiple-choice examination questions in dentistry.[30]

Dündar Sarı[16] found no significant difference between ChatGPT and Gemini in terms of overall accuracy on DUS questions; however, ChatGPT demonstrated significantly superior performance in clinical dental sciences. Notably, its higher accuracy in prosthodontics than that of Gemini indicates that model performance may vary according to the dental discipline evaluated.

When these studies are considered collectively, the discrepancies among their findings may be attributable to several factors. First, differences in question-set scope, question type, and sample size used may directly influence the performance of AI models. In addition, variations in the dental disciplines examined, differences in model versions and update timing may also contribute to differences in accuracy rates. Collectively, these discrepancies may reflect differences in dental discipline, question-set composition, sample size, model version, training data, and evaluation timing.[31]

These findings suggest that LLMs may serve as supportive tools in dental education, particularly for examination preparation, rapid access to information, and self-directed learning processes. The high accuracy rates observed especially in theoretically oriented questions indicate that LLMs may contribute to students' knowledge reinforcement and revision processes. However, several important limitations related to reliability and patient safety remain in the clinical and educational use of LLMs. Previous studies have reported that LLMs may present non-existent or unverifiable information in a fluent and convincing manner, which may lead to erroneous diagnostic or therapeutic guidance, particularly in clinical settings.[32] Evidence from dental education further indicates that LLMs may generate misinformation, provide inaccurate or fabricated references, show a limited ability to interpret visual content, and contribute to excessive dependence among students.[33] Therefore, LLM-based systems may be more appropriately used as complementary support tools supervised by educators or clinicians rather than as independent decision-making systems in dental education and clinical practice.

Although LLM-based AI models can generate correct answers, they often exhibit limitations in terms of explanatory depth, reliability of reference support, and clinical reasoning.[8] Similar concerns have been raised in studies focusing on dentistry, reporting that although LLMs may provide some correct answers to dental questions, their overall performance may remain limited, their responses may be debatable, and they may sometimes present incorrect answers with seemingly convincing reasoning.[34] Therefore, evaluating AI performance solely based on accuracy rates may not fully reflect the educational and clinical reliability of these systems.

Only the first response generated by each model was analyzed, although LLM outputs may vary across repeated queries. All questions were presented in Turkish, which may limit cross-language generalizability.[35]

Limitations

Another important limitation is the small number of questions available for each examination year. The inclusion of only 9-10 questions in each yearly subgroup may have limited the statistical power of the year-based analyses. Consequently, small changes in the number of correct responses may produce large proportional differences and distort apparent between-model differences. The implant-supported prosthodontics subgroup included only four questions, which limited the precision and generalizability of the topic-based analysis. Accordingly, these findings should be regarded as exploratory. Although generalized linear models were used in the present study, the underlying reasoning, reference support, and clinical explanations provided by the models were not analyzed in detail. Therefore, future studies should comprehensively evaluate not only accuracy rates but also the reasoning processes, source use, and clinical judgment demonstrated by AI systems. Finally, because LLMs are dynamic systems that are continuously updated, the performance results obtained in this study may be specific to the time at which the evaluation was conducted and may change with future updates.

CONCLUSION

Within the limitations of this study, the evaluated LLMs demonstrated high accuracy in answering multiple-choice prosthodontic examination questions and may have the potential to serve as complementary tools in dental education, particularly in examination preparation and information access. However, the limitations observed in scenarios requiring clinical reasoning and complex decision-making indicate that current LLMs should be regarded as supervised support systems in educational and clinical settings rather than replacements for human expertise.

ETHICAL DECLARATIONS

Ethics Committee Approval

This article does not require ethics committee approval, as it does not involve human or animal studies.

Informed Consent

Since the study was conducted without the participation of any living being, no written consent form was obtained.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The author declare no conflicts of interest related to this study.

Financial Disclosure

The author received no financial support for the conduct or publication of this research.

Author Contributions

The author is solely responsible for the entirety of conception, execution, analysis, and writing of the manuscript.

REFERENCES

1. Hoch CC, Wollenberg B, Lüers JC, et al. ChatGPT's quiz skills in different otolaryngology subspecialties: an analysis of 2576 single-choice and multiple-choice board certification preparation questions. *Eur Arch Otorhinolaryngol*. 2023;280(9):4271-4278. doi:10.1007/s00405-023-08051-4
2. Sismanoglu S, Capan BS. Performance of Artificial Intelligence on Turkish Dental Specialization Exam: can ChatGPT-4.0 and Gemini Advanced achieve comparable results to humans? *BMC Med Educ*. 2025;25(1):214. doi:10.1186/s12909-024-06389-9
3. Fogel AL, Kvedar JC. Artificial Intelligence powers digital medicine. *NPJ Digit Med*. 2018;1:5. doi:10.1038/s41746-017-0012-2
4. Cicciù M, Fiorillo L, D'Amico C, et al. 3D digital impression systems compared with traditional techniques in dentistry: a recent data systematic review. *Materials*. 2020;13(8):1982. doi:10.3390/ma13081982
5. Sezer B, Aydoğdu T. Performance of advanced Artificial Intelligence models in pulp therapy for immature permanent teeth: a comparison of ChatGPT-4 Omni, DeepSeek, and Gemini Advanced in accuracy, completeness, response time, and readability. *J Endod*. 2025;51(11):1675-1684. doi:10.1016/j.joen.2025.08.011
6. Iannantuono GM, Bracken-Clarke D, Floudas CS, Roselli M, Gulley JL, Karzai F. Applications of large language models in cancer care: current evidence and future perspectives. *Front Oncol*. 2023;13:1268915. doi:10.3389/fonc.2023.1268915
7. Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. *Future Internet*. 2023;15(9):286. doi:10.3390/fi15090286
8. Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns. *Healthcare*. 2023;11(22):2955. doi:10.3390/healthcare11222955
9. Brown T, Mann B, Ryder N, et al. Language models are few-shot learners. *Adv Neural Inf Process Syst*. 2020;33:1877-1901. doi:10.48550/arXiv.2005.14165
10. Eraslan R, Ayata M, Yagci F, Albayrak H. Exploring the potential of Artificial Intelligence chatbots in prosthodontics education. *BMC Med Educ*. 2025;25(1):321. doi:10.1186/s12909-025-06849-w
11. Dave M, Patel N. Artificial Intelligence in healthcare and education. *Br Dent J*. 2023;234(10):761-764. doi:10.1038/s41415-023-5845-2
12. Tartuk BK, Altıntaş E. Evaluation of accuracy, clinical reliability and readability of LLM-based chatbot responses in prosthetic dentistry FAQs. *Anatolian Curr Med J*. 2026;8(3):513-523. doi:10.38053/acmj.1905988
13. Çulhaoğlu AK, Kılıçarslan MA, Denız KZ. Diş Hekimliğinde Uzmanlik Sınavının farklı eğitim seviyelerdeki algı ve tercih durumlarının değerlendirilmesi. *Ata Diş Hek Fak Derg*. 2021;31(3):420-426. doi:10.17567/ataunidfd.911839
14. Tosun B, Yilmaz ZS. Comparison of Artificial Intelligence systems in answering prosthodontics questions from the dental specialty exam in Turkey. *J Dent Sci*. 2025;20(3):1454-1459. doi:10.1016/j.jds.2025.01.025
15. Venkatesan S, Krishnamoorthi D, Raju R, Mohan J, Thomas PA, Rubasree B. Evidence-based prosthodontics. *J Pharm Bioallied Sci*. 2022;14(Suppl 1):S50-S59. doi:10.4103/jpbs.jpbs_149_22
16. Dundar Sari MB, Sezer B. Comparative performance evaluation of ChatGPT-4 Omni and Gemini Advanced in the Turkish Dentistry Specialization Exam. *BMC Med Educ*. 2026;26(1):251. doi:10.1186/s12909-026-08621-0
17. Alowais SA, Alghamdi SS, Alsuhebany N, et al. Revolutionizing healthcare: the role of Artificial Intelligence in clinical practice. *BMC Med Educ*. 2023;23(1):689. doi:10.1186/s12909-023-04698-z
18. Erdal SG, Güdül A, Köroğlu A. The effectiveness of large language models in dental specialty questions: a comparative study in the field of prosthodontics. *BMC Med Educ*. 2026;26(1):455. doi:10.1186/s12909-026-08808-5
19. Yilmaz BE, Gokkurt Yilmaz BN, Ozbey F. Artificial Intelligence performance in answering multiple-choice oral pathology questions: a comparative analysis. *BMC Oral Health*. 2025;25(1):573. doi:10.1186/s12903-025-05926-2
20. Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023;2(2):e0000198. doi:10.1371/journal.pdig.0000198
21. Kasagga A, Sapkota A, Changaramkumarath G, et al. Performance of ChatGPT and large language models on medical licensing exams worldwide: a systematic review and network meta-analysis with meta-regression. *Cureus*. 2025;17(10):e94300. doi:10.7759/cureus.94300
22. Temiz M, Güzel C. Assessing the performance of ChatGPT on dentistry specialization exam questions: a comparative study with DUS examinees. *Medical Records*. 2025;7(1):162-166. doi:10.37990/medr.1567242
23. Ekici Ö. Comparative evaluation of four large language models in Turkish Dentistry Specialization Exam. *Selcuk Dent J*. 2025;12(4):6-10. doi:10.15311/selcukdentj.1674113

24. Haberal M, Hançerlioğulları D. Can Artificial Intelligence chatbots think like dentists? A comparative analysis based on dental specialty examination questions in restorative dentistry. *BMC Oral Health*. 2026; 26(1):231. doi:10.1186/s12903-025-07612-9
25. Kim J, Podlasek A, Shidara K, Liu F, Alaa A, Bernardo D. Limitations of large language models in clinical problem-solving arising from inflexible reasoning. *Sci Rep*. 2025;15(1):39426. doi:10.1038/s41598-025-22940-0
26. Karvekar S, Anand V, Singh D, Chaudhary VK, Fatima S, Mathar MI. Digital dentistry and Artificial Intelligence: a systematic review on innovations in diagnosis, treatment planning, and prosthodontics. *Cureus*. 2026;18(2):e104422. doi:10.7759/cureus.104422
27. Alfaraj A, Limones Á, Ahmad S, et al. Harnessing AI in prosthodontics and implant dentistry: an umbrella review of systematic evidence. *J Prosthodont*. 2026;35(2):127-142. doi:10.1111/jopr.70091
28. Revilla-León M, Gómez-Polo M, Vyas S, et al. Artificial Intelligence applications in implant dentistry: a systematic review. *J Prosthet Dent*. 2023;129(2):293-300. doi:10.1016/j.prosdent.2021.05.008
29. Turan Gökduman C, Arılı Öztürk E, Aktaş Ş, Çanakçi BC. Comparison of chatbots' accuracy in endodontics questions in dentistry specialization exam in Turkiye: ChatGPT-4o, Gemini Advanced, Copilot, and Claude. *BMC Oral Health*. 2025;26(1):28. doi:10.1186/s12903-025-07346-8
30. Tassoker M. ChatGPT-4 Omni's superiority in answering multiple-choice oral radiology questions. *BMC Oral Health*. 2025;25(1):173. doi:10.1186/s12903-025-05554-w
31. Çetin SG, Karadağ GG. Comparative evaluation of Artificial Intelligence models in answering pediatric dentistry questions of the Turkish Dental Specialization Exam (DUS). *Dicle Dent J*. 2025;26(2):31-37.
32. Kong M, Fok EHW, Yiu CKY. A Scoping review of large language models in dental education: applications, challenges, and prospects. *Int Dent J*. 2025;75(6):103854. doi:10.1016/j.identj.2025.103854
33. Roustan D, Bastardot F. The clinicians' guide to large language models: a general perspective with a focus on hallucinations. *Interact J Med Res*. 2025;14:e59823. doi:10.2196/59823
34. Aşık A, Kuru E. Analysis of ChatGPT's answers to pedodontics questions asked in the dentistry specialization training entrance exam: cross-sectional study. *Turkiye Klinikleri J Dental Sci*. 2025;31(3):401-406. doi:10.5336/dentalsci.2024-107488
35. Mavrych V, Ganguly P, Bolgova O. Using large language models (ChatGPT, Copilot, PaLM, Bard, and Gemini) in gross anatomy course: comparative analysis. *Clin Anat*. 2025;38(2):200-210. doi:10.1002/ca.24244

Journal of Dental Sciences and Education

Original Article

DOI: 10.51271/JDSE-0083

Evaluation of antibacterial activity of different pulp capping materials

Savaş Sağmak*[1], Ömer Çellik[2], Sadık Akgün[3]

[1]Department of Restorative Dentistry, Faculty of Dentistry, Adıyaman University, Adıyaman, Turkiye
[2]Department of Restorative Dentistry, Faculty of Dentistry, Dicle University, Diyarbakır, Turkiye
[3]Department of Medical Microbiology, Faculty of Medicine, Adıyaman University, Adıyaman, Turkiye

Cite this article as: Sağmak S, Çellik Ö, Akgün S. Evaluation of antibacterial activity of different pulp capping materials. *J Dent Sci Educ.* 2026;4(3):91-96. doi:10.51271/JDSE-0083

Received: 01.07.2026 ◆ **Accepted**: 01.08.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Aims: Pulp health is at risk from environmental factors such as caries or trauma. To prevent these risks, the pulp is treated with a capping material, either directly or indirectly. The aim of this study was to evaluate the antibacterial effects of four different pulp capping materials against *Streptococcus mutans (S. mutans)* and *Lactobacillus acidophilus (L. acidophilus)*.

Methods: The antimicrobial activity of four different pulp capping materials (TheraCal LC, Dycal, MTA Angelus, and MTA Cem LC) and a chlorhexidine control group was assessed using the agar diffusion method at three-time intervals (1, 3, and 5 days). For each experimental group, twenty samples were prepared and tested against *S. mutans* (ATCC® 25175) and *L. acidophilus* (ATCC® 11975), which served as standard bacterial strains. The obtained data were analyzed using repeated-measures ANOVA followed by Bonferroni post hoc tests for multiple comparisons.

Results: Based on the overall mean values across all measurement periods, the chlorhexidine control group exhibited the highest antimicrobial activity against both *S. mutans* and *L. acidophilus*. Analyzing the experimental groups, MTA Angelus, TheraCal LC, and Dycal showed antimicrobial activity against *S. mutans*. MTA Cem LC demonstrated no discernible antimicrobial activity. However, TheraCal LC and Dycal demonstrated antimicrobial activity against *L. acidophilus*. MTA Angelus and MTA Cem LC exhibited no antimicrobial activity. Statistically significant differences were observed among the groups on all measurement days ($p<0.05$).

Conclusion: The antibacterial performance of pulp capping materials was influenced by both material composition and bacterial species. These findings highlight the importance of considering antibacterial properties alongside biological characteristics when selecting pulp capping materials.

Keywords: Pulp capping, antibacterial activity, *Streptococcus mutans*, *Lactobacillus acidophilus*, mineral trioxide aggregate

INTRODUCTION

Vital pulp therapy aims to preserve the vitality of pulp tissues compromised by deep carious lesions, restorative procedures, or traumatic injuries in the absence of pulp necrosis. This therapeutic approach, commonly referred to as pulp capping, involves the placement of a biocompatible protective material over the exposed pulp to promote tertiary dentin formation.[1]

Accurate assessment of pulp status prior to the capping procedure is critical for determining treatment prognosis. Clinical diagnosis is guided by responses to thermal stimuli, pain on percussion, and radiographic findings.[2] In vital pulp therapy, adequate pulp perfusion and periodontal health are fundamental to preserving pulp vitality. Effective hemorrhage control, thorough disinfection of the cavity walls, and the selection of an appropriate capping material are equally essential for achieving favorable treatment outcomes.[3]

Pulp capping is indicated in cases of deep carious lesions or when pulp exposure results from carious progression. Dental caries arise from the destruction of hard tooth tissues by acidogenic bacteria; therefore, reducing bacterial load and activity is pivotal in arresting carious progression and protecting the pulp. *Streptococcus mutans (S. mutans)* and *Lactobacillus acidophilus (L. acidophilus)* are the principal microorganisms implicated in caries development.[4] *S. mutans*, which plays a central role in caries initiation, is recognized as the predominant cariogenic species in the oral flora. Its presence across various caries types is attributed to its acidogenic and aciduric characteristics.[5] *L. acidophilus* is a gram-positive bacterium that metabolizes sugars to produce lactic acid. Although frequently identified in active carious lesions, it is not considered a primary initiator of caries due to its relatively late colonization; nonetheless, it is implicated in the progression of dentinal caries in advanced stages.[6]

Vital pulp therapy pursues two principal objectives: arresting carious progression and preventing its reactivation. This begins with the elimination of communication between the carious cavity and the oral environment. Following rubber dam isolation, infected carious tissue is removed and bacterial

*Corresponding Author: Savaş Sağmak, savassagmak@adiyaman.edu.tr

contamination is reduced. The subsequent restoration should incorporate antibacterial materials capable of inhibiting bacterial proliferation and metabolic activity.[7]

Materials intended for pulp capping must fulfill a range of essential criteria, including the capacity to stimulate dentin repair, exert antibacterial effects, promote pulp healing, demonstrate biocompatibility, and restrict lesion advancement.[8] In recent years, the development of novel vital pulp treatment materials has generated considerable research interest regarding their antimicrobial properties.

Among the capping materials evaluated in the present study, Dycal is a calcium hydroxide-based cavity liner with well-established clinical use. Its principal properties include acid neutralization, antibacterial activity, stimulation of secondary dentin formation, radiopacity, and ease of clinical application.[9,10] TheraCal LC is a resin-modified, light-cured, calcium silicate-containing pulp capping material suitable for both direct and indirect pulp capping. It provides a protective barrier for the pulp and is composed of calcium oxide, strontium, silica, barium sulfate, calcium silicate, barium zirconate, Bis-GMA, and polydimethacrylate. Its composition consists of approximately 45% mineral content, 10% radiopaque filler, 5% hydrophilic thickener, and 40-45% resin. TheraCal LC is notable for its sustained release of calcium ions.[11] MTA Angelus is a mineral trioxide aggregate-based material whose principal constituents include tricalcium silicate, tricalcium aluminate, tricalcium oxide, and silicon oxide. Its indications encompass root canal furcation repair, management of internal resorption, pulpotomy, and both direct and indirect pulp capping.[12] MTA Cem LC is a light-polymerized, resin-modified, tricalcium silicate-based material indicated for use as a protective liner beneath light-cured base materials and as a co-lining agent in deep cavities or cases involving pulp exposure.[13]

Antibacterial activity is assessed through various methodologies, including disk and well agar diffusion tests and dilution assays. Among these, the agar diffusion technique is widely regarded as the method of choice owing to its cost-effectiveness, technical simplicity, and capacity for the simultaneous comparison of multiple test materials.[14]

Although the antibacterial properties of several pulp capping materials have been investigated, limited information is available regarding newly introduced resin-modified calcium silicate-based materials. In particular, the antibacterial activity of MTA Cem LC has not been previously reported. Therefore, the present study aimed to evaluate the antibacterial efficacy of four pulp capping materials-TheraCal LC, MTA Angelus, Dycal, and MTA Cem LC-against *S. mutans* and *L. acidophilus*, two microorganisms actively involved in carious progression and pulp inflammation. The null hypothesis of the study was as follows: there is no statistically significant difference in antibacterial activity against *S. mutans* and *L. acidophilus* among the tested pulp capping materials.

METHODS

Ethics

Since the study was conducted exclusively under laboratory conditions using previously available bacterial strains, Ethics Committee approval was not required in accordance with the applicable institutional and national regulations. All procedures were carried out in accordance with the ethical rules and the principles.

The present study is an in vitro microbiological laboratory study. The bacterial strains used in this research were obtained from the laboratory collection of the Public Health Institution of Turkiye. No human participants, patient samples, clinical specimens, personal data, or animal subjects were involved in the study.

Study Design

This study has been completed at the Faculty of Dentistry, Adıyaman University. The laboratory phase of the study was completed over a three-month period. The aim of the study was to evaluate the antibacterial activity of four experimental groups and one control group against *S. mutans* (ATCC® 25175) and *L.* acidophilus (ATCC® 11975) on days 1, 3, and 5.

Preparation of Samples

An a priori power analysis was conducted using G*Power 3.1.9.7 for a repeated-measures ANOVA (within-between interaction). Based on a large effect size [f(V)=0.50], α=0.05, and 95% power, the minimum required sample size was 50 specimens. Accordingly, 50 specimens were analyzed for *S. mutans* and 50 for *L. acidophilus* (n=10 per group), resulting in a total of 100 specimens. For each experimental group, twenty specimens measuring 4 mm in depth and 7 mm in diameter were fabricated using cylindrical plastic molds (**Table 1**). Specimens in groups 1 and 4 were polymerized using a blue light-emitting LED for the durations specified by the manufacturers. In contrast, specimens in groups 2 and 3 were prepared by manual mixing on a glass slab with a metal spatula and subsequently placed into the molds. All samples were covered with a mylar strip and glass coverslips and allowed to set at 25°C. All procedures during sample preparation were performed under sterile conditions using sterile instruments, and the specimens were aseptically placed onto the appropriate culture media without risk of contamination.

Table 1. Experimental groups, manufacturers and material contents

Material	Manufacturer	Components	LOT number
TheraCal LC	Bisco, Lançon De Provence, France	**Light-curing single paste:** Resin (Bis-GMA and PEGD) modified calcium silicate, filled with CaO, calcium silicate particles (type III Portland cement), Sr glass, fumed silica, barium sulfate, and barium zirconate	2300110466
Dycal	Dentsply Tulsa Dental, Johnson City, TN, USA	Two-paste system made of a base paste (1,3-butylene glycol disalicylate, zinc oxide, calcium phosphate, calcium tungstate, and iron oxide pigments) and a catalyst paste (calcium hydroxide, N-ethyl-o/p-toluene sulphonamide, zinc oxide, titanium oxide, zinc stearate, and iron oxide pigments)	023304
MTA Angelus	Angelus, Londrina, Brazil	**Powder:** tri-calcium silicate, di-calcium silicate, tri-calcium aluminate, ferroaluminate, tri-calcium, calcium oxide, bismuth oxide **Liquid:** distilled water	102273
MTA Cem LC	GC, Tokyo, Japan	BisGMA, CaO, silica, barium sulphate (BaSO4)	LC210714

LC: Light-cured, Bis-GMA: Bisphenol A-glycidyl methacrylate, PEGD: Polyethylene glycol dimethacrylate, CaO: Calcium oxide, Sr: Strontium, $BaSO_4$: Barium sulfate

Materials tested:

- **Group 1:** TheraCal LC
- **Group 2:** Dycal
- **Group 3:** MTA Angelus
- **Group 4:** MTA Cem LC
- **Group 5 (Control):** Chlorhexidine (2%, BEST, Turkiye)

Assessment of Antibacterial Activity

For each experimental group, a total of 20 samples were prepared, including 10 samples (n=10) inoculated with *S. mutans* (ATCC® 25175) and 10 samples (n=10) inoculated with *L. acidophilus* (ATCC® 11975) standard bacterial isolates.[11] For the agar diffusion test, sheep blood agar (SBA) (ThermoScientific Oxford, UK) and MRS (Man, Ragosa & Sharpe) media were used. Lyophilized *S. mutans* (ATCC® 25175) and *L. acidophilus* (ATCC® 11975) strains were cultured in appropriate growth media and incubated at 37°C under CO_2-enriched conditions. Following incubation, bacterial suspensions were prepared in physiological saline and adjusted to a 0.5 McFarland turbidity standard (1.5×10^8 CFU/ml). A 100 µL aliquot of the standardized suspension was uniformly spread onto the surface of agar plates using sterile cotton swabs. Wells measuring 7 mm in diameter were created in the agar plates using the blunt end of a micropipette tip, and the wells were completely filled with the test materials (**Figure 1, 2**).

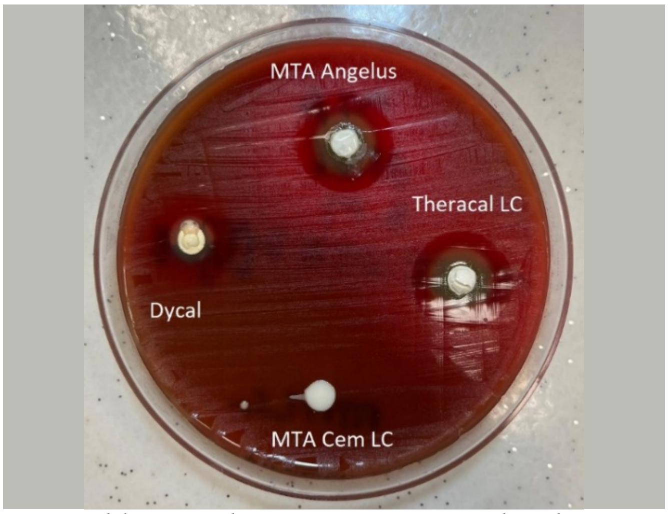

Figure 1. Inhibition zone diameters against *S. mutans* on the 1st day

S. mutans: Streptococcus mutans, MTA: Mineral trioxide aggregate, LC: Light-cured, CEM: Calcium-enriched mixture

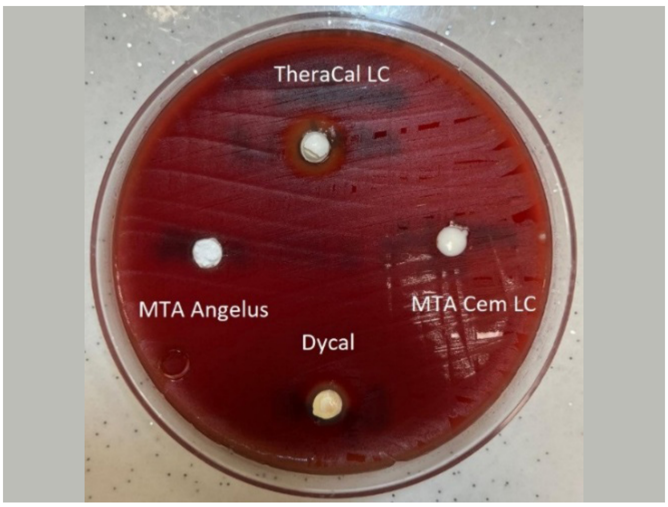

Figure 2. Inhibition zone diameters against *L. acidophilus* on the 1st day

L. acidophilus: Lactobacillus acidophilus, LC: Light-cured, MTA: Mineral trioxide aggregate, CEM: Calcium-enriched mixture

Statistical Analysis

All data analysis of the data obtained in this study was performed using SPSS version 28.0 (IBM Corp., Armonk, NY, USA). The normality of data distribution was assessed using the Shapiro-Wilk test. Results were expressed as mean±standard deviation (SD). Changes over time were evaluated using repeated measures analysis of variance (repeated measures ANOVA). Bonferroni post hoc tests were used for multiple comparisons. $p<0.05$ was considered significant.

RESULTS

The antibacterial activity of the tested pulp capping materials against *S. mutans* is presented in **Figure 3**. When the mean values of the three measurement periods were evaluated, the chlorhexidine-treated control group demonstrated the highest antimicrobial activity against *S. mutans*. Among the experimental materials, MTA Angelus showed the greatest antibacterial effect, followed by TheraCal LC and Dycal, respectively. A reduction in antimicrobial effectiveness was observed in subsequent measurements for these groups. In contrast, MTA Cem LC did not exhibit any detectable antimicrobial activity. Statistically significant differences among the groups were identified on all measurement days ($p<0.05$) (**Table 2**).

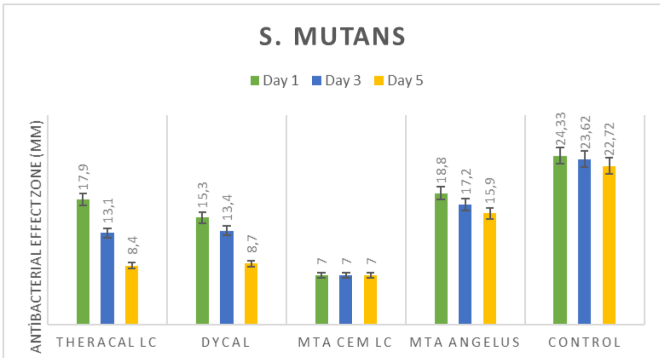

Figure 3. Graphical representation of the effectiveness obtained from the experimental groups against *S. mutans*

S. mutans: Streptococcus mutans, MTA: Mineral trioxide aggregate, CEM: Calcium-enriched mixture, LC: Light-cured

Table 2. Antimicrobial activity values against *S. mutans*

Groups	Day 1	Day 3	Day 5	Mean/SD
TheraCal LC	17.90 ± 0.73^c	13.10 ± 0.99^c	8.40 ± 0.69^c	13.13 ± 1.33^c
Dycal	15.30 ± 1.15^c	13.40 ± 1.34^c	8.70 ± 0.94^c	12.47 ± 0.74^c
MTA Angelus	18.80 ± 1.03^b	17.20 ± 0.63^b	15.90 ± 0.73^b	17.30 ± 0.74^b
MTA Cem LC	7.00 ± 0.00^d	7.00 ± 0.00^d	7.00 ± 0.00^d	7.00 ± 0.00^d
Chlorhexidine	24.33 ± 0.37^a	23.62 ± 0.42^a	22.72 ± 0.66^a	23.60 ± 0.43^a

Values are presented as mean±SD. Different superscript letters indicate statistically significant differences between groups according to Bonferroni post hoc comparisons ($p<0.05$). *S. mutans: Streptococcus mutans*, SD: Standard deviation, LC: Light-cured, MTA: Mineral trioxide aggregate, CEM: Calcium-enriched mixture

The data obtained from the experimental groups against *L. acidophilus* are presented in **Figure 4**. Based on the mean values obtained from three measurement periods, the chlorhexidine-treated control group exhibited the highest antimicrobial activity against *L. acidophilus*. Among the experimental materials, TheraCal LC and Dycal demonstrated antibacterial effects following the control group; however, a reduction in their efficacy was observed in subsequent measurements. In contrast, neither MTA Angelus nor MTA Cem LC showed any detectable antimicrobial activity against

L. acidophilus. Statistically significant differences among all groups were identified at each measurement interval ($p<0.05$) (**Table 3**).

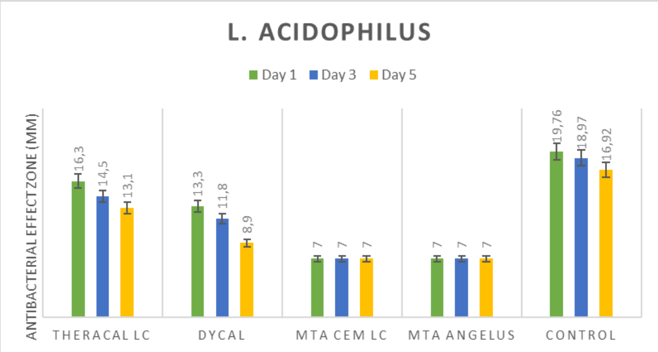

Figure 4. Graphical representation of the data obtained from the experimental groups against *L. acidophilus*

L. acidophilus: Lactobacillus acidophilus, MTA: Mineral trioxide aggregate, CEM: Calcium-enriched mixture, LC: Light-cured

Table 3. Antimicrobial activity values against *L. acidophilus*

Groups	Day 1	Day 3	Day 5	Mean/SD
TheraCal LC	16.30±0.67^{b}	14.50±0.53^{b}	13.10±0.57^{b}	14.63±0.43^{b}
Dycal	13.30±0.95^{c}	11.80±0.92^{c}	8.90±1.20^{c}	11.33±0.90^{c}
MTA Angelus	7.00±0.00^{d}	7.00±0.00^{d}	7.00±0.00^{d}	7.00±0.00^{d}
MTA Cem LC	7.00±0.00^{d}	7.00±0.00^{d}	7.00±0.00^{a}	7.00±0.00^{d}
Chlorhexidine	19.76±0.85^{a}	18.87±0.55^{a}	16.92±1.40^{a}	18.86±0.79^{a}

Values are presented as mean±SD. Different superscript letters indicate statistically significant differences between groups according to Bonferroni post hoc comparisons ($p<0.05$). *L. acidophilus: Lactobacillus acidophilus*, SD: Standard deviation, LC: Light-cured, MTA: Mineral trioxide aggregate

DISCUSSION

Microorganisms play a central role in the etiology of dental caries. The dominant bacteria and their metabolic byproducts in deep dentinal caries that have reached the pulp pose a direct threat to pulp health. Preserving pulp vitality therefore requires the use of a biocompatible material with proven antibacterial properties.[15] Accordingly, *S. mutans* the predominant species in enamel caries and *L. acidophilus* the dominant species in dentinal caries were selected as the test microorganisms in this study.

The agar disk diffusion method is widely employed in dental research owing to its technical simplicity and cost-effectiveness; however, a recognized limitation of this approach is its inability to characterize the specific mechanisms or temporal dynamics of the antimicrobial effect.[14]

Pulp exposure may arise as a complication of prosthetic treatment, trauma, or carious progression. The primary objective of vital pulp therapy is to preserve the health of the pulp tissue. Both direct and indirect pulp capping protect pulp vitality by establishing a biological barrier against thermal, chemical, and microbial insults.[16] The essential function of a capping material is to shield the exposed pulp and sustain its vitality through the stimulation of reparative dentin formation. Calcium hydroxide, a cornerstone of vital pulp therapy, induces a superficial necrotic zone on the exposed pulp surface, thereby triggering hard tissue deposition through repair mechanisms. Simultaneously, calcium hydroxide applied over a thin residual dentin layer undergoes dissolution, facilitating the release of growth factors from the pulp.[17] In addition to organic solvent- or water-based suspensions and calcium salicylate-based paste formulations such as Dycal, resin monomer-based calcium hydroxide preparations are also available. Beyond these conventional materials, contemporary bioactive pulp capping agents have been developed. Bioactive materials are defined as therapeutic agents designed to initiate biomineralization while inherently possessing antibacterial properties.[18] Tricalcium silicate-based materials such as MTA and Biodentine do not contain calcium oxide in their composition; however, calcium hydroxide is generated upon contact with tissue fluids.[19] In the present study, four widely used pulp capping materials were evaluated; TheraCal LC, Dycal, MTA Angelus, and MTA Cem LC.

TheraCal LC is a light-cured, resin-modified, calcium silicate-based pulp capping material. The release of hydroxyl ions (OH^{-}) from TheraCal LC establishes a highly alkaline environment with bactericidal properties.[20] Poggio et al.[21] assessed the antibacterial effects of six pulp capping materials-Dycal, Calcicur, Calcimol LC, TheraCal LC, MTA Angelus, and Biodentine-against three streptococcal species and reported that TheraCal LC demonstrated antibacterial activity against *S. mutans*. Similarly, Fathy et al.[22] investigated the antimicrobial effects of Biodentine and TheraCal LC against *S. mutans* and found that, although Biodentine exhibited greater antibacterial efficacy, the difference between the two materials was not statistically significant. Our findings are consistent with previous reports demonstrating the antibacterial activity of TheraCal LC against *S. mutans*.

Gandolfi et al.[23] reported that TheraCal LC releases higher levels of calcium than ProRoot MTA and Dycal, yet displays considerably lower solubility, while the pH of TheraCal LC was comparable to that of ProRoot MTA. Overall, TheraCal LC has been shown to exhibit antibacterial activity comparable to that of calcium hydroxide and calcium silicate-based materials.[21]

Dycal is a calcium hydroxide-based material used for both direct and indirect pulp capping that sets via a chemical reaction. Upon dissolution in water, it dissociates into calcium and hydroxyl ions. The resulting alkaline environment disrupts bacterial cell wall integrity and induces protein denaturation, thereby conferring antibacterial properties.[10] Moradi et al.,[24] in a comparative study involving Fuji II LC glass ionomer, Dycal, Calcimol LC, TheraCal LC, and ACTIVA BioACTIVE, reported that Dycal exhibited the highest antimicrobial activity among all tested groups.

The elevated pH of MTA results from the sustained release of calcium ions and the subsequent formation of calcium hydroxide. This high alkalinity, combined with significant ion release, underlies the antibacterial properties attributed to MTA.[25] While MTA has demonstrated efficacy against certain facultative bacteria, it has been shown to be ineffective against anaerobic species. Kim et al.[26] found that MTA Angelus exhibited greater antibacterial activity against *S. mutans* compared to two other mineral trioxide aggregate formulations, Endocem MTA and ProRoot MTA.

Elshamy et al.[25] compared the antimicrobial properties of MTA, Dycal, and bioceramic materials against *S. mutans* and *L. acidophilus* using an agar diffusion test, and determined that MTA exhibited the highest level of antimicrobial activity. The authors attributed this finding to the strong correlation between MTA's elevated pH and its pronounced antibacterial properties.

MTA Cem LC is a recently introduced light-cured, resin-modified, tricalcium silicate-based material indicated for direct and indirect pulp capping as well as for use as a protective base or liner beneath composite restorations, amalgams, cements, and other restorative base materials. To date, no studies evaluating the antibacterial efficacy of this material have been reported in the literature. In the present study, MTA Cem LC did not demonstrate any inhibitory effect on the tested bacteria. This may be attributed to the reduction in ion release associated with its resin content, which likely impairs the degree of pH elevation and consequently diminishes its antimicrobial potential. The absence of published data on this material underscores the need for further investigation. Considering its resin-modified formulation, future studies should evaluate not only its antibacterial activity but also calcium ion release, alkalinity, and bioactivity using complementary in vitro and in vivo models.

Bakir et al.[27] reported that MTA demonstrated antibacterial activity against *S. mutans, L. acidophilus and Enterococcus faecalis (E. faecalis)*. Zarrabi et al.[28] similarly reported antibacterial activity of MTA against *E. faecalis* and *S. mutans*.

Poggio et al.[21] conducted a comparative analysis of the antibacterial activity of six pulp dressing materials against *Streptococcus salivarius*, *S. sanguis*, and *S. mutans*. Among the materials tested, MTA Angelus, Calcicur, Dycal, and TheraCal LC demonstrated inhibitory activity against *S. mutans*, with MTA Angelus producing the most pronounced effect, followed by TheraCal LC, Dycal, and Calcicur, in descending order. Biodentine and Calcimol LC did not exhibit any measurable antimicrobial activity.[29] The findings of the present study are in agreement with these results; MTA Angelus was the most effective material against *S. mutans*, followed by TheraCal LC and Dycal, respectively. MTA Cem LC, however, did not demonstrate any antimicrobial effect.

The antibacterial activity of pulp capping materials may be considered a relevant factor in the clinical decision-making process during vital pulp therapy. Although the success of vital pulp therapy is primarily associated with the preservation of pulp vitality and the induction of reparative dentin formation, the presence of residual microorganisms following caries removal remains a potential challenge. Therefore, materials exhibiting antibacterial properties may provide a biological advantage by limiting bacterial activity at the pulp-dentin interface and supporting a more favorable environment for pulpal healing.[15] However, the clinical performance of pulp capping materials is determined by a combination of factors, and antibacterial activity should be evaluated together with other important characteristics, such as biocompatibility, bioactivity, sealing ability, and mechanical properties.[30]

Limitations

This study has several limitations. First, the antibacterial activity of the tested materials was evaluated only under in vitro conditions using the agar diffusion method, which may not fully reflect the complex biological environment of the oral cavity. Moreover, this method primarily evaluates the inhibitory effect of materials and may be influenced by their diffusion properties within the agar medium. Second, only two cariogenic bacterial species were investigated, whereas clinical dental caries is associated with a more diverse microbial community. In addition, the present study focused on antibacterial activity and did not evaluate other biological properties of the materials, such as biocompatibility, bioactivity, or long-term interaction with pulp tissues. Therefore, further in vivo and long-term clinical studies involving broader microbial models and comprehensive biological evaluations are needed to confirm the clinical relevance of the present findings.

CONCLUSION

Within the limitations of this in vitro study, the antibacterial activity of pulp capping materials was found to vary according to both the material composition and the bacterial species tested. While MTA Angelus, TheraCal LC, and Dycal exhibited antibacterial activity against *S. mutans*, only TheraCal LC and Dycal demonstrated activity against *L. acidophilus*. In contrast, MTA Cem LC showed no detectable antibacterial effect against either bacterial species under the experimental conditions. These findings suggest that the antimicrobial performance of pulp capping materials may be influenced by their composition and should be considered alongside their biological and physicochemical properties when selecting a material for vital pulp therapy. Further well-designed in vivo and long-term clinical studies are required to confirm the clinical relevance of these findings.

ETHICAL DECLARATIONS

Ethics Committee Approval

Since the study was conducted exclusively under laboratory conditions using previously available bacterial strains, Ethics Committee approval was not required in accordance with the applicable institutional and national regulations.

Informed Consent

Since the study was conducted without the participation of any living being, no written consent form was obtained.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: SS, ÖC; Design: SS, ÖC; Control: SS, ÖC; Resources: SS, SA; Materials: SS, SA; Data Collection and/or Processing: SS, SA; Analysis and/or Interpretation: SS, SA; Literature Review: SS, ÖC, SA; Writing the Article: SS; Critical Review: SS, ÖC, SA.

REFERENCES

1. Hilton TJ. Keys to clinical success with pulp capping: a review of the literature. *Oper Dent.* 2009;34(5):615-625. doi:10.2341/09-132-0
2. Arandi NZ. Calcium hydroxide liners: a literature review. *Clin Cosmet Investig Dent.* 2017;9:67-72. doi:10.2147/CCIDE.S141381

3. Kunert M, Lukomska-Szymanska M. Bio-inductive materials in direct and indirect pulp capping-a review article. *Materials (Basel).* 2020;13(5): 1204. doi:10.3390/ma13051204
4. Ghoddusi J, Forghani M, Parisay I. New approaches in vital pulp therapy in permanent teeth. *Iran Endod J.* 2014;9(1):15-22.
5. Ge Y, Caufield PW, Fisch GS, Li Y. *Streptococcus mutans* and *Streptococcus sanguinis* colonization correlated with caries experience in children. Caries Res. 2008;42(6):444-448. doi:10.1159/000159608
6. Beighton D. The complex oral microflora of high-risk individuals and groups and its role in the caries process. *Community Dent Oral Epidemiol.* 2005;33(4):248-255. doi:10.1111/j.1600-0528.2005.00232.x
7. Mickenautsch S, Yengopal V, Banerjee A. Pulp response to resin-modified glass ionomer and calcium hydroxide cements in deep cavities: a quantitative systematic review. *Dent Mater.* 2010;26(8):761-770. doi:10.1016/j.dental.2010.03.021
8. Song M, Kim S, Kim T, et al. Development of a direct pulp-capping model for the evaluation of pulpal wound healing and reparative dentin formation in mice. *J Vis Exp.* 2017;(119):54973. doi:10.3791/54973
9. Furey A, Hjelmhaug J, Lobner D. Toxicity of flow line, Durafill VS, and Dycal to dental pulp cells: effects of growth factors. *J Endod.* 2010;36(7): 1149-1153. doi:10.1016/j.joen.2010.03.013
10. Gandolfi MG, Siboni F, Botero T, Bossù M, Riccitiello F, Prati C. Calcium silicate and calcium hydroxide materials for pulp capping: biointeractivity, porosity, solubility and bioactivity of current formulations. *J Appl Biomater Funct Mater.* 2015;13(1):43-60. doi:10.5301/jabfm.5000201
11. Hebling J, Lessa FC, Nogueira I, Carvalho RM, Costa CA. Cytotoxicity of resin-based light-cured liners. *Am J Dent.* 2009;22(3):137-142.
12. Moussa, SA. Mineral trioxide aggregate (MTA) vs. calcium hydroxide in direct pulp capping literature review. *Online J Dent Oral Heal.* 2018; 1:1-6.
13. Ipek I, Karaağaç Eskibağlar B, Yildiz Ş, Ataş O, Ünal M. Analysis of the bond strength between conventional, putty or resin-modified calcium silicate cement and bulk fill composites. *Aust Dent J.* 2023;68(4):265-272. doi:10.1111/adj.12977
14. Maalej SM, Meziou MR, Rhimi FM, Hammami A. Comparison of disc diffusion, Etest and agar dilution for susceptibility testing of colistin against *Enterobacteriaceae*. *Lett Appl Microbiol.* 2011;53(5):546-551. doi:10.1111/j.1472-765X.2011.03145.x
15. Bhavana V, Chaitanya KP, Gandi P, Patil J, Dola B, Reddy RB. Evaluation of antibacterial and antifungal activity of new calcium-based cement (Biodentine) compared to MTA and glass ionomer cement. *J Conserv Dent.* 2015;18(1):44-46. doi:10.4103/0972-0707.148892
16. Aguilar P, Linsuwanont P. Vital pulp therapy in vital permanent teeth with cariously exposed pulp: a systematic review. *J Endod.* 2011;37(5):581-587. doi:10.1016/j.joen.2010.12.004
17. Leye Benoist F, Gaye Ndiaye F, Kane AW, Benoist HM, Farge P. Evaluation of mineral trioxide aggregate (MTA) versus calcium hydroxide cement [Dycal(®)] in the formation of a dentine bridge: a randomised controlled trial. *Int Dent J.* 2012;62(1):33-39. doi:10.1111/j.1875-595X.2011.00084.x
18. Schmalz G, Widbiller M, Galler KM. Material tissue interaction--from toxicity to tissue regeneration. *Oper Dent.* 2016;41(2):117-131. doi:10.2341/15-249-BL
19. Wadia R. Biodentine vs MTA. *Br Dent J.* 2025;238(4):247. doi:10.1038/s41415-025-8458-0
20. Takita T, Hayashi M, Takeichi O, et al. Effect of mineral trioxide aggregate on proliferation of cultured human dental pulp cells. *Int Endod J.* 2006;39(5):415-422. doi:10.1111/j.1365-2591.2006.01097.x
21. Poggio C, Arciola CR, Beltrami R, et al. Cytocompatibility and antibacterial properties of capping materials. *Scientific World J.* 2014; 2014:181945. doi:10.1155/2014/181945
22. Fathy SM, Abd El-Aziz AM, Labah DA. Cellular interaction and antibacterial efficacy of two hydraulic calcium silicate-based cements: Cell-dependent model. *J Conserv Dent.* 2019;22(1):17-22. doi:10.4103/JCD.JCD_272_18
23. Gandolfi MG, Siboni F, Prati C. Chemical-physical properties of TheraCal, a novel light-curable MTA-like material for pulp capping. *Int Endod J.* 2012;45(6):571-579. doi:10.1111/j.1365-2591.2012.02013.x
24. Moradi Z, Abbasi M, Bahador A, et al. Evaluation of cytotoxicity and antibacterial activity of different pulp capping liners. *Biomater Investig Dent.* 2023;10(1):2287019. doi:10.1080/26415275.2023.2287019
25. Elshamy FM, Singh G, Elraih H, Gupta I, Idris FA. Antibacterial effect of new bioceramic pulp capping material on the main cariogenic bacteria. *J Contemp Dent Pract.* 2016;17(5):349-353. doi:10.5005/jp-journals-10024-1854
26. Kim RJ, Kim MO, Lee KS, Lee DY, Shin JH. An in vitro evaluation of the antibacterial properties of three mineral trioxide aggregate (MTA) against five oral bacteria. *Arch Oral Biol.* 2015;60(10):1497-1502. doi:10.1016/j.archoralbio.2015.07.014
27. Bakır EP, Bakır Ş, Ünal S. Comparison of antibacterial effects of pulp capping materials. *Selcuk Dent J.* 2021;8(2):553-560. doi:10.15311/selcukdentj. 896007
28. Hasan Zarrabi M, Javidi M, Naderinasab M, Gharechahi M. Comparative evaluation of antimicrobial activity of three cements: new endodontic cement (NEC), mineral trioxide aggregate (MTA) and Portland. *J Oral Sci.* 2009;51(3):437-442. doi:10.2334/josnusd.51.437
29. Poggio C, Beltrami R, Colombo M, Ceci M, Dagna A, Chiesa M. In vitro antibacterial activity of different pulp capping materials. *J Clin Exp Dent.* 2015;7(5):e584-8. doi:10.4317/jced.52401
30. Hanna SN, Perez Alfayate R, Prichard J. Vital pulp therapy an insight over the available literature and future expectations. *Eur Endod J.* 2020; 5(1):46-53. doi:10.14744/eej.2019.44154

Journal of Dental Sciences and Education

Original Article

DOI: 10.51271/JDSE-0084

Learning preferences, distance education perceptions, and anxiety (STAI-S-6) among dentistry students

Luísa Schubach da Costa Barreto*, Maria Cardoso de Castro Berry, Flavia Artese, Cátia Cardoso Abdo Quintão

Department of Social and Preventive Dentistry (PRECOM), Faculty of Dentistry, Rio de Janeiro State University, Rio de Janeiro, Brazil

Cite this article as: Barreto LSC, Berry MCC, Artese F, Quintão CCA. Learning preferences, distance education perceptions, and anxiety (STAI-S-6) among dentistry students. *J Dent Sci Educ.* 2026;4(3):97-104. doi:10.51271/JDSE-0084

Received: 13.01.2026 ◆ **Accepted**: 13.08.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Aims: To investigate undergraduate and graduate dentistry students' perceptions of distance education (DE), learning preferences, and anxiety levels during the early phase of the COVID-19 pandemic, and to compare these outcomes according to academic level.

Methods: A multicenter cross-sectional study was conducted between March and June 2020 using a 44-item questionnaire organized into five domains. Anxiety was assessed using the Brazilian Portuguese version of the State-Trait Anxiety Inventory Short Form (STAI-S-6). Comparisons between undergraduate and graduate students were performed using Chi-square tests, independent samples t-tests, and one-way analysis of variance, adopting a significance level of 5%.

Results: A total of 293 students participated (150 undergraduates and 143 graduate students). Undergraduate students were significantly younger and more frequently belonged to lower socioeconomic strata ($p<0.001$). Female students, undergraduates, and participants from intermediate-to-lower socioeconomic strata (particularly class C1 according to the Brazilian Economic Classification Criteria) exhibited higher anxiety levels ($p<0.05$). A weak negative correlation was observed between age and anxiety scores ($r=-0.125$; $p=0.032$). Significant differences between groups were observed regarding perceptions of DE, learning preferences, academic performance, and the perceived relationship between DE and anxiety ($p<0.001$). Students reporting extremely negative effects of DE on academic performance and those perceiving that DE greatly increased anxiety presented the highest STAI-S-6 scores ($p<0.001$).

Conclusion: These findings suggest that dentistry students experienced considerable psychological burden during the early phase of the COVID-19 pandemic. Undergraduate students appeared to report greater difficulties adapting to DE and higher anxiety levels than graduate students. The results may indicate the importance of mental-health-centered educational strategies and equitable access to digital resources to support students during periods of educational disruption.

Keywords: Dentistry, perception, distance education, COVID-19, anxiety

INTRODUCTION

The COVID-19 pandemic abruptly disrupted higher education worldwide, compelling institutions to adopt emergency distance education (DE) with little time for pedagogical restructuring.[1] This sudden transition reshaped academic routines and affected students' psychological well-being,[2] with meta-analytic evidence indicating a high prevalence of anxiety, depression, and stress among students receiving DE during this period.[3,4] Although vaccination campaigns and the return to face-to-face activities have occurred, more recent studies have described educational and psychological challenges associated with remote and hybrid learning environments. These observations are presented only to provide context for understanding the experiences reported during the early stages of the COVID-19 pandemic.[1,3,5-8]

In dentistry, these challenges were amplified by the central role of clinical and preclinical training.[9] The sudden interruption of laboratory and clinical activities, concerns regarding infection, and uncertainty about academic progression were associated with emotional distress among dental students.[3,4,7] A recent meta-analysis estimated that nearly one-third of dental students experienced anxiety symptoms during the pandemic,[6] illustrating the magnitude of the psychological burden observed in this population.

Early studies conducted during the first waves of COVID-19 reported higher levels of anxiety, stress, and depressive symptoms, particularly among younger students, female, and those experiencing difficulties adapting to DE.[10] More recent investigations have described similar findings, indicating that DE, social isolation, and reduced clinical exposure were associated with less favorable emotional outcomes.[11,12] Although more recent studies have documented persistent academic fatigue and anxiety associated with remote and

*Corresponding Author: Luísa Schubach da Costa Barreto, luisaschubach@gmail.com

hybrid learning environments, these observations are presented only to contextualize findings obtained during the early phase of the COVID-19 pandemic.[1,13,14]

Perceptions of impaired academic performance, lower content quality, and greater difficulty maintaining concentration appear to be particularly relevant, yet their relationship with validated anxiety measures remains insufficiently explored. To assess psychological impact with methodological rigor, the State-Trait Anxiety Inventory (STAI) remains one of the most frequently used instruments worldwide,[15] including in educational and clinical settings. The Brazilian short-form version of the state scale (STAI-S-6) demonstrates robust validity and reliability and is suitable for large online surveys requiring efficient and low-burden assessment of anxiety symptoms.[16]

Despite the growing body of literature examining the educational and psychological consequences of the COVID-19 pandemic, relatively few studies have directly compared undergraduate and graduate dentistry students. Because these groups differ in age, academic responsibilities, professional expectations, and levels of autonomy, they may have experienced emergency DE differently during the early phase of the pandemic. Moreover, most previous investigations have focused primarily on undergraduate students, with limited information regarding possible differences in learning preferences, perceptions of DE, and anxiety levels between academic stages. A better understanding of these differences may provide valuable insights for the development of more resilient, equitable, and student-centered educational strategies.

Therefore, the aim of this multicenter cross-sectional study was to investigate learning preferences, perceptions of DE, and state anxiety (STAI-S-6) among undergraduate and graduate dentistry students during the early phase of the COVID-19 pandemic, and to compare these outcomes according to academic level. In addition, the associations between sociodemographic characteristics, DE experiences, and anxiety scores were explored. It was hypothesized that undergraduate and graduate students would differ in their perceptions of DE and anxiety profiles.

METHODS

Ethics

The study protocol was approved by the This study was approved by the Universitário Pedro Ernesto Ethics Committee (Date: 17.05.2020, Decision No: 4.029.405, CAAE No: 30637420.3.0000.0008), and all participants provided electronic informed consent before enrollment. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Study Design

This cross-sectional observational study was designed and reported in accordance with the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) guidelines[17] for cross-sectional studies and adhered to the CHERRIES (Checklist for Reporting Results of Internet E-Surveys)[18] criteria to ensure transparency and methodological rigor during online data collection.

This multicenter cross-sectional study was conducted between March and June 2020, during the early phase of the COVID-19 pandemic, using a 44-item online questionnaire organized into five domains.

Participants, Eligibility Criteria, and Recruitment

Participants included undergraduate and graduate dentistry students enrolled in multiple dental schools in Brazil during the period of emergency DE. Recruitment was conducted through institutional communication channels, including e-mail, WhatsApp Messenger, and Facebook academic groups, in order to maximize geographic and institutional representation. Eligibility criteria included current enrollment in undergraduate or graduate dental programs and voluntary agreement to participate in the study. Questionnaires with incomplete responses were excluded from the analysis.

Survey Administration and Structure

Data were collected using a self-administered questionnaire created in Google Forms and organized into five domains: (1) demographic and socioeconomic characteristics; (2) social isolation and concerns related to the COVID-19 pandemic; (3) preferences regarding distance education; (4) perceptions of the impact of DE on academic performance and anxiety; and (5) psychological assessment using the Brazilian Portuguese version of the State-Trait Anxiety Inventory Short Form (STAI-S-6) (**Table 1**).[16]

To prevent duplicate submissions and ensure the integrity of the dataset, participants were required to sign in to a Google account before accessing the survey. This verification procedure was used exclusively to restrict responses to one submission per participant and did not allow investigators to identify respondents. No personally identifiable information was collected, and all responses were analyzed anonymously. Access to the database was restricted to the research team and protected by password-secured files, ensuring confidentiality and compliance with ethical standards. Socioeconomic status was classified according to the Brazilian Economic Classification Criteria (ABEP), which categorizes households into strata ranging from A (highest) to E (lowest) based on household assets and educational level.

Anxiety Assessment

Psychological assessment was performed using the validated Brazilian Portuguese version of the STAI-S-6, which consists of six items derived from the original 20-item state anxiety scale.16 Responses were recorded on a four-point Likert scale ranging from 1 (“not at all”) to 4 (“very much”). Positively worded items were reverse scored, and the resulting raw scores were converted to the standardized STAI metric according to the procedure proposed by Fioravanti-Bastos et al.[16] Higher scores corresponded to higher levels of state anxiety.

Sample Size Calculation

Sample size calculations were performed using G*Power software (version 3.1; Heinrich-Heine University, Düsseldorf, Germany). Sample size estimation was based on detecting a clinically meaningful difference in STAI-S-6 anxiety scores between undergraduate and graduate students. Preliminary estimates of variability were obtained from a pilot study conducted with dentistry students, which yielded standard deviations of 11.41 and 7.97 for undergraduate and graduate students, respectively. A minimum difference of 6.67 points was derived from the distribution of pilot scores and was

Table 1. Schematic representation of the questions included in the online questionnaire to collect information on undergraduate and graduate dentistry students

Section 1. Sociodemographic data (10 questions) **Questions about your personal life, status and academic degree**
Q1. Gender: ( ) Male ( ) Female
Q2. How old are you?
Q3. What race do you declare yourself? ( ) White ( ) Black ( ) Brown ( ) Yellow ( ) Indigenous
Q4. Which religion do you stand for? ( ) Catholic ( ) Protestant ( ) Spiritism ( ) Jewish ( ) Buddhist ( ) No religion to state
Q5. What is your civil status? ( ) Single ( ) Married ( ) Divorced ( ) Widower
Q6. Mark the Brazilian region you currently live in: ( ) North ( ) Northeast ( ) Southeast ( ) Midwest ( ) South
Q7. How many people currently live in the household with you? ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q8. Where are you taking virtual dentistry classes? ( ) Graduation ( ) Postgraduate courses
Q9. Which educational institution are you attending? ( ) Public ( ) Private
Q10. What is your highest academic degree reached? ( ) Complete high school ( ) Graduate ( ) Specialist ( ) Masters ( ) Doctorate
Section 2. Socioeconomic criteria (15 questions) **Questions about items in the household that you currently live for economic classification. All items of electronics that are listed below must be working, including those that are stored. If they are not working, consider only if you intend to repair or replace them within the next 6 months**
Q11. Number of passenger cars exclusively for private use: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q12. Number of monthly employees, considering only those who work at least 5 days a week: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q13. Number of washing machines, excluding six pack: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q14. Number of bathrooms: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q15. Quantity of DVD, including any device that reads DVD and disregards car DVD: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q16. Number of refrigerators: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q17. Number of independent freezers or part of duplex refrigerator: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q18. Number of microcomputers, considering desktop computers, laptops, notebooks and netbooks, disregarding tablets, palms or smartphones ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q19. Quantity of dishwasher: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q20. Quantity of microwave oven: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q21. Number of motorcycles, excluding those used exclusively for professional use: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q22. Number of clothes drying machines, considering lava and drought: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q23. The water used at home comes from: ( ) General distribution network ( ) Well or spring ( ) Other source
Q24. Considering the street segment of your household, you would say that the street is: ( ) None ( ) 1 ( ) 2 ( ) 3 ( ) 4 or more
Q25. How educated is the head of the family? Consider as the head of the household the person who contributes most of the household income in question ( ) Illiterate/elementary 1 incomplete ( ) Elementary 1 completed/elementary 2 incomplete ( ) Completed elementary/high school incomplete ( ) Completed high school/incomplete higher education ( ) Completed higher education
Section 3. Work relationships (2 questions) **For each question, the participant can only choose one option**
Q26. What is your main concern regarding the coronavirus pandemic? ( ) Professional instability, the chance of unemployment ( ) Economic crisis in the country ( ) Contagion ( ) Food ( ) Education
Q27. How are you respecting the quarantine during social isolation? ( ) I do not leave home for nothing ( ) I am staying home as much as possible (going out only to buy food/medicine) ( ) I am going out as usual
Section 4. Methods of distance education (11 questions) **For each question, the participant can only choose one option**
Q28. During COVID-19 pandemic, what is your frequency of virtual lessons? ( ) Always ( ) Many times ( ) Sometimes ( ) Rarely ( ) Never
Q29. Which of the following options would be your preferred distance education modality? ( ) Video lessons recorded and available for me to watch for as long as I want, without the possibility of asking questions ( ) Video lessons recorded and available for me to watch for as long as I want, with the possibility of taking questions by e-mail ( ) Unrecorded video lessons, only with the teacher's virtual presence, without sharing slides on the screen, without the student's face screen and without the possibility of asking questions. ( ) Video lessons not recorded, with virtual presence of the teacher, with slide sharing on the screen, with the screen of the student's face and with the possibility of asking questions ( ) Unrecorded video lessons, without the teacher's virtual presence, only with his live voice, with sharing slides on the screen, without the screen of the student's face and with the possibility of answering questions.
Q30. Did you have difficulty attending online classes? ( ) No difficulty ( ) Low difficulty ( ) Average difficulty ( ) A lot of difficulty ( ) Extreme difficulty
Q31. What was the biggest difficulty encountered when attending classes online? ( ) Not applicable (as theoretically reported having no difficulty in the previous question) ( ) Impaired Wi-Fi connection ( ) How to make questions ( ) Lack of interaction between student and teacher ( ) I can't focus ( ) Platform adaptation
Q32. What would be your suggestion for improving distance learning methods? ( ) Standardization and organization of content with creation of schedules and goals to develop student autonomy ( ) Improve dynamics, have a summary at the end of the class, provide complementary material such as PDF documents and exercises ( ) Availability to answer questions with an online tutor ( ) Video storage lessons so I can watch as many times as I want ( ) Improving the stability of videoconferencing platforms
Q33. Regarding the online class hours, what time do you define as ideal for one lesson? ( ) 15 minutes ( ) 30 minutes ( ) 45 minutes ( ) 1 hour ( ) More than 1 hour
Q34. What is your class modality of preference? ( ) Online class ( ) Face-to-face class
Q35. Mark the option that justifies the chosen preference: ( ) More organized content ( ) Practicality in answering questions ( ) Greater interactivity between student and teacher ( ) I stay more concentrated ( ) I am used to it
Q36. Video lessons require the student to look at the computer screen for a long period of time. How do you think it affects your performance during distance education? ( ) Affects extremely negatively ( ) Affects negatively ( ) Does not affect ( ) Affects positively ( ) Affects extremely positively
Q37. Regarding the content of the course, how do you evaluate online classes? ( ) Very limited ( ) Limits a little ( ) No effect ( ) Covers a little ( ) Completely covers
Q38. To what extent do you believe that distance learning activities affect students' anxiety levels? ( ) Increases a lot ( ) Increases a little ( ) Does not affect ( ) Reduces a little ( ) Reduces a lot
Section 5. Psychological assessment (6 questions) **Your personal self-assessment regarding the CURRENT moment. Please select one of the numbers 1, 2, 3 or 4 to indicate how you feel NOW. There are no right or wrong answers. Do not spend too much time on any of the statements. Try to mark an answer that comes closest to how you feel IN THIS MOMENT. For each question, the participant can only choose one option**
Q39. I feel calm: ( ) 1=not at all ( ) 2=somewhat ( ) 3=moderately ( ) 4=very much
Q40. I am tense: ( ) 1=not at all ( ) 2=somewhat ( ) 3=moderately ( ) 4=very much
Q41. I feel at easy: ( ) 1=not at all ( ) 2=somewhat ( ) 3=moderately ( ) 4=very much
Q42. I feel nervous: ( ) 1=not at all ( ) 2=somewhat ( ) 3=moderately ( ) 4=very much
Q43. I am relaxed: ( ) 1=not at all ( ) 2=somewhat ( ) 3=moderately ( ) 4=very much
Q44. I am worried: ( ) 1=not at all ( ) 2=somewhat ( ) 3=moderately ( ) 4=very much

Q: Question, COVID-19: Coronavirus disease 2019, Wi-Fi: Wireless fidelity, PDF: Portable document format, DVD: Digital versatile disc, AI: Artificial Intelligence, STAI: State-Trait Anxiety Inventory considered clinically meaningful. Assuming a two-sided significance level of 5% and a statistical power of 90%, intentionally chosen to reduce the probability of type II error, a minimum sample of 92 participants (46 per group) was required. The corresponding expected standardized effect size (Cohen's d) was approximately 0.68, indicating a moderate-to-large effect.[20]

Statistical Analysis

All data analyses were performed using SPSS software (IBM SPSS Statistics, version 21.1; IBM Corp., Armonk, NY, USA). Descriptive statistics were calculated for all variables. Data distribution was assessed using the Shapiro-Wilk test.

Comparisons between undergraduate and graduate students were performed using independent samples t-tests for continuous variables and Chi-square tests for categorical variables. One-way analysis of variance (ANOVA) was used to compare mean STAI-S-6 scores across categories containing more than two groups. When appropriate, post hoc comparisons were performed, and statistically significant differences between groups were indicated by superscript letters in the tables.

Pearson's correlation coefficient was used to evaluate the association between age and STAI-S-6 scores. Correlation coefficients were interpreted according to conventional thresholds, with values below 0.30 considered weak, between 0.30 and 0.50 moderate, and above 0.50 strong. Statistical significance was established at $p<0.05$.

RESULTS

Sample Characterization

A total of 293 dentistry students completed the survey, including 150 undergraduate and 143 graduate students. The mean age of participants was 26.5 years, with undergraduate students being significantly younger than graduate students (22.0±3.4 vs. 31.0±7.4 years; $p<0.001$; Cohen's d=1.58). Most respondents were female (82.9%), enrolled in public institutions (67.6%), and belonged to economic classes A or B (78.9%) according to ABEP criteria. The STAI-S-6 demonstrated satisfactory internal consistency in the present sample, with a Cronbach's alpha coefficient of 0.82. A weak but statistically significant negative correlation was observed between age and STAI-S-6 scores ($r=-0.125$; $p=0.032$) (**Figure 1**). Significant differences between undergraduate and graduate students were observed across several demographic and socioeconomic variables (**Table 2**).

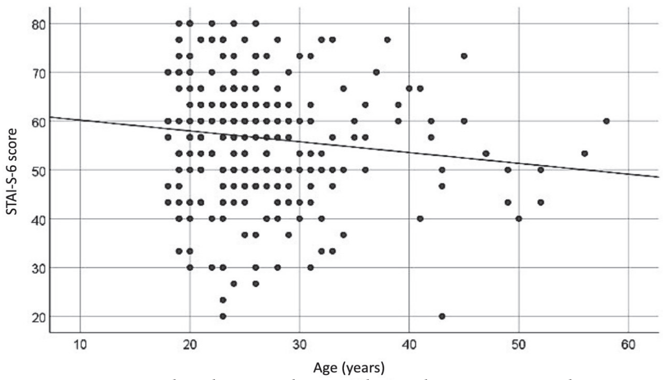

Figure 1. Scatter plot showing the correlation between age and STAI-S-6 anxiety scores among dentistry students. The fitted regression line illustrates the weak negative association identified ($r=-0.125$; $p=0.032$)

STAI-S-6: State-Trait Anxiety Inventory-State 6-item short form

Table 2. Profile and between-group comparison of demographic, socioeconomic, and academic variables in the study sample

	Groups			
	Total (n=293) n (%)	Undergraduate (n=150) n (%)	Graduate (n=143) n (%)	p*
Gender				
Female	243 (82.9)	124 (82.7)	119 (83.2)	>0.999
Male	50 (17.1)	26 (17.3)	24 (16.8)	
Race				
White	193 (65.9)	86 (57.3)	107 (74.8)	0.002
Black	22 (7.5)	18 (12)	4 (2.8)	
Brown	74 (25.3)	43 (28.7)	31 (21.7)	
Others	4 (1.4)	3 (2)	1 (0.7)	
Religion				
Catholic	129 (44.3)	51 (34.5)	78 (54.5)	0.014
Protestant	51 (17.5)	31 (20.9)	20 (14)	
Spiritism	21 (7.2)	11 (7.4)	10 (7)	
Buddhist	3 (1)	2 (1.4)	1 (0.7)	
No religion to state	87 (29.9)	53 (35.8)	34 (23.8)	
Civil status				
Single	231 (78.8)	144 (96)	87 (60.8)	<0.001
Married	57 (19.5)	6 (4)	51 (35.7)	
Divorced	5 (1.7)	0 (0)	5 (3.5)	
Brazilian region				
North	18 (6.1)	7 (4.7)	11 (7.7)	0.066
Northeast	49 (16.7)	33 (22)	16 (11.2)	
Southeast	209 (71.3)	103 (68.7)	106 (74.1)	
South	17 (5.8)	7 (4.7)	10 (7)	
Number of people currently living in the household				
None	18 (6.1)	5 (3.3)	13 (9.1)	0.002
1	49 (16.7)	15 (10)	34 (23.8)	
2	71 (24.2)	39 (26)	32 (22.4)	
3	78 (26.6)	44 (29.3)	34 (23.8)	
4 or more	77 (26.3)	47 (31.3)	30 (21)	
Institution				
Public	198 (67.6)	93 (62)	105 (73.4)	0.050
Private	95 (32.4)	57 (38)	38 (26.6)	
Highest academic degree				
Completed high school	148 (50.5)	148 (98.7)	0 (0)	<0.001
Graduated, DDS	47 (16)	2 (1.3)	45 (31.5)	
Specialist	39 (13.3)	0 (0)	39 (27.3)	
Masters, MSD	53 (18.1)	0 (0)	53 (37.1)	
Doctorate, PhD	6 (2)	0 (0)	6 (4.2)	
Economic class rate				
A	87 (29.7)	35 (23.3)	52 (36.4)	<0.001
B1	65 (22.2)	23 (15.3)	42 (29.4)	
B2	79 (27)	46 (30.7)	33 (23.1)	
C1	44 (15)	33 (22)	11 (7.7)	
C2	18 (6.1)	13 (8.7)	5 (3.5)	
	Mean (SD)	**Mean (SD)**	**Mean (SD)**	**p****
Age (years)	26 (7.1)	22 (3.4)	31 (7.4)	<0.001

Legend: *Chi-square test for categorical variables and **independent samples t-test for age. DDS: Doctor of dental surgery, MSD: Master of science in dentistry, PhD: Doctor of philosophy, SD: Standard deviation

Comparison Between Groups

Significant differences between undergraduate and graduate students were observed regarding social isolation, pandemic-related concerns, preferences related to DE, perceived difficulties, instructional modality, perceived academic performance, content quality, and the perceived relationship between DE and anxiety (**Table 3**; $p<0.05$).

Table 3. Comparison of perceptions, concerns, and DE-related variables between undergraduate and graduate dentistry students

	Total n (%)	Groups: Undergraduate (n=150) n (%)	Groups: Graduate (n=143) n (%)	p*
Social isolation				
I am going out as usual	36 (12.3)	5 (3.3)	31 (21.7)	<0.001
I am staying home as much as possible (going out only to buy food/medicine)	195 (66.6)	95 (63.3)	100 (69.9)	
I do not leave home for nothing	62 (21.2)	50 (33.3)	12 (8.4)	
Concern				
Professional instability, the chance of unemployment	28 (9.6)	9 (6)	19 (13.3)	0.001
Economic crisis in the country	70 (23.9)	27 (18)	43 (30.1)	
Contagion	181 (61.8)	102 (68)	79 (55.2)	
Food	4 (1.4)	4 (2.7)	0 (0)	
Education	10 (3.4)	8 (5.3)	2 (1.4)	
Frequency of virtual lessons				
Many times	37 (12.6)	14 (9.3)	23 (16.1)	0.083
Sometimes	62 (21.2)	30 (20)	32 (22.4)	
Rarely	115 (39.2)	57 (38)	58 (40.6)	
Never	79 (27)	49 (32.7)	30 (21)	
Distance education preferences				
Video lessons recorded and available for me to watch for as long as I want, without the possibility of asking questions	5 (1.7)	5 (3.3)	0 (0)	0.002
Video lessons recorded and available for me to watch for as long as I want, with the possibility of taking questions by email	172 (58.7)	95 (63.3)	77 (53.8)	
Unrecorded video lessons, only with the teacher's virtual presence, without sharing slides on the screen, without the student's face screen and without the possibility of asking questions	3 (1)	0 (0)	3 (2.1)	
Video lessons not recorded, with virtual presence of the teacher, with slide sharing on the screen, with the screen of the student's face and with the possibility of asking questions	85 (29)	33 (22)	52 (36.4)	
Unrecorded video lessons, without the teacher's virtual presence, only with his live voice, with sharing slides on the screen, without the screen of the student's face and with the possibility of answering questions	28 (9.6)	17 (11.3)	11 (7.7)	
Difficulty				
No difficulty	59 (20.1)	17 (11.3)	42 (29.4)	<0.001
Low difficulty	79 (27)	41 (27.3)	38 (26.6)	
Average difficulty	97 (33.1)	51 (34)	46 (32.2)	
A lot of difficulty	38 (13)	25 (16.7)	13 (9.1)	
Extreme difficulty	20 (6.8)	16 (10.7)	4 (2.8)	
Difficulty justification				
Not applicable (as theoretically reported having no difficulty in the previous question)	47 (16)	18 (12)	29 (20.3)	0.348
Impaired Wi-Fi connection	72 (24.6)	41 (27.3)	31 (21.7)	
How to make questions	7 (2.4)	5 (3.3)	2 (1.4)	
Lack of interaction between student and teacher	19 (6.5)	11 (7.3)	8 (5.6)	
I can't focus	126 (43)	64 (42.7)	62 (43.4)	
Platform adaptation	22 (7.5)	11 (7.3)	11 (7.7)	
Improvement suggestion				
Standardization and organization of content with creation of schedules and goals to develop student autonomy	56 (19.1)	23 (15.3)	33 (23.1)	0.142
Improve dynamics, have a summary at the end of the class, provide complementary material such as PDF documents and exercise	124 (42.3)	74 (49.3)	50 (35)	
Availability to answer questions with an online tutor	10 (3.4)	5 (3.3)	5 (3.5)	
Video storage lessons so I can watch as many times as I want	82 (28)	39 (26)	43 (30.1)	
Improving the stability of videoconferencing platforms	21 (7.2)	9 (6)	12 (8.4)	
Ideal time of lesson				
<30 minutes	41 (14)	20 (13.4)	21 (14.7)	0.709
45 minutes	92 (31.5)	45 (30.2)	47 (32.9)	
1 hour	118 (40.4)	65 (43.6)	53 (37.1)	
More than 1 hour	41 (14)	19 (12.8)	22 (15.4)	
Class modality				
Online class	22 (7.5)	3 (2)	19 (13.3)	<0.001
Face-to-face class	271 (92.5)	147 (98)	124 (86.7)	
Class modality justification				
More organized content	12 (4.1)	4 (2.7)	8 (5.6)	0.451
Practicality in answering questions	11 (3.8)	6 (4)	5 (3.5)	
Greater interactivity between student and teacher	78 (26.6)	38 (25.3)	40 (28)	
I stay more concentrated	157 (53.6)	80 (53.3)	77 (53.8)	
I am used to it	35 (11.9)	22 (14.7)	13 (9.1)	
Performance				
Affects extremely negatively	44 (15)	31 (20.7)	13 (9.1)	<0.001
Affects negatively	141 (48.1)	80 (53.3)	61 (42.7)	
Does not affect	108 (36.9)	39 (26)	69 (48.3)	
Content				
Very limited	71 (24.2)	55 (36.7)	16 (11.2)	<0.001
Limits a little	116 (39.6)	61 (40.7)	55 (38.5)	
No effect	43 (14.7)	14 (9.3)	29 (20.3)	
Covers a little	37 (12.6)	14 (9.3)	23 (16.1)	
Completely covers	26 (8.9)	6 (4)	20 (14)	
Relationship DE/anxiety				
Increases a lot	90 (30.7)	65 (43.3)	25 (17.5)	<0.001
Increases a little	125 (42.7)	58 (38.7)	67 (46.9)	
Does not affect	39 (13.3)	16 (10.7)	23 (16.1)	
Reduces a little	39 (13.3)	11 (7.3)	28 (19.6)	
Anxiety (categorized)				
Has no anxiety	16 (5.5)	9 (6)	7 (4.9)	0.874
Higher anxiety	277 (94.5)	141 (94)	136 (95.1)	

Legend: *Chi-square test. Underlined categories indicate proportions with standardized adjusted residuals ≥1.96. Wi-Fi: Wireless fidelity, PDF: Portable document format, DE: Distance education

Anxiety State Scale Impact

STAI-S-6 scores differed significantly according to gender, academic level, socioeconomic status, perceived academic performance, and the perceived relationship between DE and anxiety (**Table 4; Figures 2 and 3**).

Table 4. Between-group comparisons of mean STAI-S-6 scores across demographic, educational, and behavioral variables

	STAI-S-6 score (SD)	p
Gender[1]		
Female	57.4 (12.5)	0.008
Male	52.3 (12.4)	
Group[1]		
Undergraduate	58.4 (13.0)	0.009
Graduate	54.6 (11.9)	
Institution[1]		
Public	56.4 (12.8)	0.803
Private	56.8 (12.4)	
Economic class rate[2]		
A	53.9^{a} (12.9)	0.022
B1	56.9^{ab} (10.7)	
B2	55.7^{ab} (12.8)	
C1	60.3^{b} (11.6)	
C2/D/E	62.0^{ab} (16.4)	
Social isolation[2]		
I am going out as usual	55.6 (10.5)	0.730
I am staying home as much as possible (going out only to buy food/medicine)	56.4 (12.8)	
I do not leave home for nothing	57.6 (13.1)	
Frequency of virtual lessons[2]		
Always/many times	55.2 (13.8)	0.226
Sometimes	54.0 (12.9)	
Rarely	57.7 (11.5)	
Never	57.6 (13.2)	
Ideal time of lesson[2]		
<30 minutes	55.0 (13.0)	0.104
45 minutes	58.9 (12.5)	
1 hour	56.1 (12.0)	
More than 1 hour	53.7 (13.6)	
Class modality[1]		
Online class	51.7 (10.9)	0.059
Face-to-face class	56.9 (12.7)	
Performance[2]		
Affects extremely negatively	61.1^{a} (13.1)	<0.001
Affects negatively	58.1^{ab} (11.7)	
Does not affect/affects positively or extremely positively	52.7^{b} (12.6)	
Self-perception-relationship DE/anxiety[2]		
Increases a lot	62.7^{a} (11.9)	<0.001
Increases a little	54.2^{b} (10.7)	
Does not affect	54.4^{b} (12.4)	
Reduces a little/too much	52.1^{b} (15.2)	

Legend: 1-Independent samples t-test; 2-One-way ANOVA. Means with different superscript letters indicate statistically significant differences between groups (p<0.05). Means sharing at least one letter are not significantly different. STAI-S-6: State-Trait Anxiety Inventory-State 6-item short form, SD: Standard deviation, DE: Distance education, ANOVA: Analysis of variance

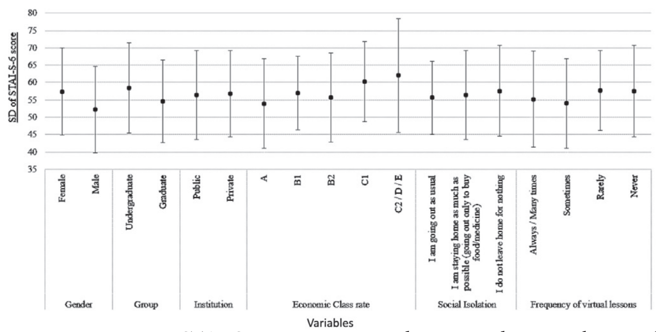

Figure 2. Mean STAI-S-6 scores according to demographic and socioeconomic characteristics. Black dots represent mean values and vertical bars indicate±standard deviation

STAI-S-6: State-Trait Anxiety Inventory-State 6-item short form

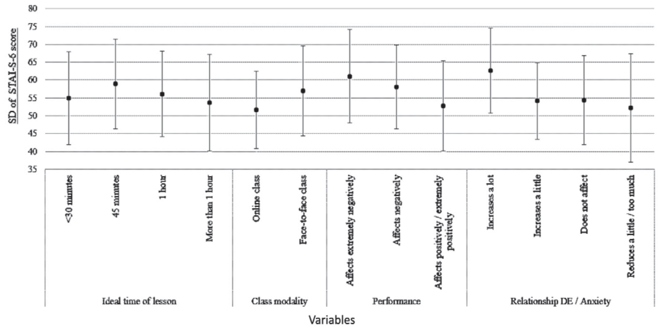

Figure 3. Mean STAI-S-6 anxiety scores and standard deviations according to DE-related variables. Black dots represent the mean STAI-S-6 scores for each DE category, and vertical bars indicate±standard deviation

STAI-S-6: State-Trait Anxiety Inventory-State 6-item short form, DE: Distance education

Students reporting extremely negative effects of DE on academic performance presented the highest STAI-S-6 scores (61.1±13.1), whereas those reporting no negative effects exhibited lower scores (52.7±12.6; p<0.001; η^2=0.11).

Similarly, students who perceived that DE greatly increased anxiety exhibited the highest STAI-S-6 scores (62.7±12.4), significantly higher than those observed in the remaining categories (p<0.001; η^2=0.18). No significant differences in STAI-S-6 scores were observed according to social isolation status, frequency of virtual classes, or preferred class duration (p>0.05).

To complement the statistical significance tests, effect size estimates were calculated for all significant comparisons involving STAI-S-6 scores (**Table 5**). A large effect size was observed for age differences between undergraduate and graduate students (Cohen's d=1.58). Small-to-moderate effects were identified for gender (Cohen's d=0.41) and academic level (Cohen's d=0.31). For variables analyzed using one-way ANOVA, the largest effect size was found for the perceived relationship between DE and anxiety (η^2=0.18), followed by the perceived impact of DE on academic performance (η^2=0.11), whereas socioeconomic status showed a smaller effect (η^2=0.05).

Table 5. Effect sizes for significant comparisons involving STAI-S-6 scores

Variable	Statistical test	Effect size	Magnitude
Age (undergraduate vs. graduate students)	Independent samples t-test	Cohen's d=1.58	Large
Gender	Independent samples t-test	Cohen's d=0.41	Small-to-moderate
Academic level	Independent samples t-test	Cohen's d=0.31	Small
Socioeconomic status	One-way ANOVA	η^2=0.05	Small
Perceived impact of DE on academic performance	One-way ANOVA	η^2=0.11	Moderate
Perceived relationship between DE and anxiety	One-way ANOVA	η^2=0.18	Large

Legend: Interpretation of effect sizes followed conventional criteria. For Cohen's d, values of approximately 0.20, 0.50, and 0.80 were considered small, moderate, and large, respectively. For eta squared (η^2), values of approximately 0.01, 0.06, and 0.14 were interpreted as small, moderate, and large effects, respectively. STAI-S-6: State-Trait Anxiety Inventory-State 6-item short form, DE: Distance education, ANOVA: Analysis of variance

DISCUSSION

This study provides insight into dentistry students' psychological and academic experiences during the early phase of the COVID-19 pandemic, when emergency DE was abruptly implemented. Relatively few studies have directly compared undergraduate and graduate students, and the present findings suggest that academic level may influence how students perceive DE and its relationship with anxiety.

More recent studies have reported similar educational and psychological challenges in hybrid and DE settings, although caution is warranted when comparing these findings with data obtained during the early phase of the pandemic.[11,12,14,21]

An important finding of the present study was that subjective perceptions related to DE, particularly perceived academic performance decline and the belief that DE was associated with increased anxiety, were accompanied by higher STAI-S-6 scores. These findings suggest that students' appraisal of DE may have been more strongly related to psychological distress than structural characteristics of DE itself.[22-24]

Differences in social isolation patterns and pandemic-related concerns were observed between undergraduate and graduate students. Graduate students appeared to be more concerned about professional instability and the economic consequences of the pandemic, whereas undergraduate students more frequently reported concerns related to contagion and educational issues. These differences may reflect distinct academic responsibilities and career stages, findings that are consistent with previous reports describing heterogeneous psychological responses among health professions students.[24,25]

The interruption of clinical and preclinical activities was associated with emotional distress among dental students, consistent with previous studies showing that reduced opportunities for practical training, uncertainty regarding academic progression, and concerns about professional preparedness are associated with anxiety and perceived academic impairment.[3,4,9,26] Limited teacher-student interaction and reduced opportunities for hands-on learning have also been identified as stressors that may negatively affect student engagement and perceived quality of education.[7,10,13]

Undergraduate students exhibited higher anxiety levels than graduate students. This finding may reflect differences in academic maturity, self-regulation skills, familiarity with independent learning, and expectations regarding professional training. Similar patterns have been reported in studies evaluating emotional responses across different educational stages.[10,11,27] Socioeconomic disparities also emerged as an important factor, with students from lower economic strata presenting higher anxiety scores, which is consistent with evidence suggesting greater academic and digital burdens among socially vulnerable students.[28] Although the correlation between age and STAI-S-6 scores reached statistical significance, its magnitude was weak ($r=-0.125$), indicating limited practical relevance. Therefore, age alone should not be considered a major determinant of anxiety levels among dentistry students.

A noteworthy contribution of this study was the association between subjective DE experiences, including perceived academic performance decline and self-perceived anxiety related to DE, and objective anxiety scores measured by the STAI-S-6. Similar relationships have been described in studies reporting that cognitive overload, prolonged screen exposure, and perceived reductions in learning quality are associated with less favorable psychological outcomes.[10-12,21,27] Recent investigations have also suggested that DE fatigue may be accompanied by reduced concentration and increased frustration.[10,14,28]

The findings of the present study underscore the importance of pedagogical approaches that combine flexibility with structured and interactive educational activities. Strategies such as shorter synchronous sessions, active learning opportunities, and improved instructional design may help reduce cognitive burden and enhance student engagement.[10-12,22] In dental education, combining digital resources with supervised practical activities and hybrid learning models may contribute to improving academic continuity and student well-being.

Although Artificial Intelligence (AI)-based educational tools were not investigated in the present study, recent technological advances may provide opportunities to address some of the educational challenges identified during emergency DE. Future studies are warranted to evaluate the potential role of AI-assisted educational approaches in dental education.

Limitations

Several limitations should be considered when interpreting the present findings. First, the cross-sectional design precludes causal inference. Second, the use of self-reported measures may have introduced reporting bias. Third, because the survey was administered online, selection bias cannot be excluded. Furthermore, data collection was conducted between March and June 2020, and therefore the findings reflect the specific circumstances of the early phase of the COVID-19 pandemic. Consequently, caution is warranted when extrapolating these results to later phases of the pandemic or contemporary educational settings.

Potentially relevant confounding variables, including previous psychiatric disorders, psychotropic medication use, social support, and academic performance before the pandemic, were not assessed and may have influenced anxiety scores. Although multivariable regression models may be useful for identifying independent predictors of anxiety, such analyses were beyond the exploratory objectives of the present study. Future longitudinal investigations are warranted to better elucidate the relative contribution of these factors to psychological outcomes among dentistry students.

CONCLUSION

These findings suggest that dentistry students experienced considerable psychological burden during emergency DE implemented in the early phase of the COVID-19 pandemic, with higher anxiety levels observed among females, undergraduate students, and individuals from lower socioeconomic backgrounds. The results may indicate that difficulties adapting to DE and the perception that DE increased anxiety were associated with higher STAI-S-6 scores. Overall, these findings underscore the potential importance of mental-health-oriented educational strategies and equitable access to digital resources during periods of major educational disruption. Lessons learned from the early pandemic period may help inform the development of more resilient and student-centered educational models in the future.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study was approved by the Universitário Pedro Ernesto Ethics Committee (Date: 17.05.2020, Decision No: 4.029.405).

Informed Consent

Electronic informed consent was obtained from all individual participants prior to their inclusion in the study. Participants were fully informed about the study's aims, procedures, potential risks and benefits, and their rights, including the right to withdraw at any time without consequence. All participants voluntarily provided electronic informed consent.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)-Finance Code 001 and the Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ)-Finance Code E-26/203.049/2023.

Author Contributions

Concept: LSCB, FRGA, CCAQ; Design: LSCB, FRGA, CCAQ; Control: LSCB, FRGA, CCAQ; Resources: LSCB, FRGA, CCAQ; Materials: LSCB; Data Collection and/or Processing: LSCB; Analysis and/or Interpretation: LSCB, MCCB; Literature Review: LSCB, MCCB, FRGA, CCAQ; Writing the Article: LSCB, MCCB, FRGA, CCAQ; Critical Review: LSCB, MCCB, FRGA, CCAQ.

Acknowledgements

The authors express their sincere gratitude to all dentistry students who generously participated in this study by responding to the online questionnaire.

REFERENCES

1. Jehi T, Khan R, Dos Santos H, Majzoub N. Effect of COVID-19 outbreak on anxiety among students of higher education: a review of literature. *Curr Psychol.* 2022;42(20):17475-17489. doi:10.1007/s12144-021-02587-6
2. Xu T, Wang H. High prevalence of anxiety, depression, and stress among remote learning students during the COVID-19 pandemic: evidence from a meta-analysis. *Front Psychol.* 2023;13:1103925. doi:10.3389/fpsyg.2022.1103925
3. Limeira FIR, Cruz PF, Andrade ARC, Guiomarino Neto AB, Arantes DC. Depression, anxiety and stress among dental students during COVID-19 pandemic and distance learning. *Rev ABENO.* 2022;22(2):1-12. doi:10.30979/revabeno.v22i2.1563
4. Silva PGB, Oliveira CAL, Borges MMF, et al. Distance learning during social seclusion by COVID-19: improving the quality of life of undergraduate dentistry students. *Eur J Dent Educ.* 2021;25(1):124-134. doi:10.1111/eje.12583
5. Kim HR, Kim EJ. Factors associated with mental health among international students during the COVID-19 pandemic in South Korea. *Int J Environ Res Public Health.* 2021;18(21):11381. doi:10.3390/ijerph182111381
6. Xie H, Wang J, Chen Y. Prevalence of anxiety and depression symptoms among international students during the COVID-19 pandemic: a meta-analysis. *INPLASY.* 2022. doi:10.37766/inplasy2022.10.0012
7. Keskin G. Self-report measurement of depression, anxiety, and stress caused by COVID-19 pandemic in senior undergraduate dental students. *Pesqui Bras Odontopediatria Clín Integr.* 2021;21:102. doi:10.1590/pboci.2021.102
8. Andersson MA, Hitlin S. Subjective dignity and self-reported health: results from the United States before and during the COVID-19 pandemic. *SSM Ment Health.* 2022;2:100113. doi:10.1016/j.ssmmh.2022.100113
9. Demirekin ZB, Buyukcavus MH. Effect of distance learning on quality of life, anxiety and stress levels of dental students during the COVID-19 pandemic. *BMC Med Educ.* 2022;22(1):3382. doi:10.1186/s12909-022-03382-y
10. Amboy MKQ, Habaña MP, Ramirez LM, Sotelo MG, Penuela AC, Oducado RMF. Zoom fatigue and satisfaction with online learning among Filipino nursing students. *Innoeduca Int J Technol Educ Innov.* 2023;9(1):53-66. doi:10.24310/innoeduca.2023.v9i1.15402
11. Zanini GS, Campanhã LMF, Biazus EL, Cardoso HF, Verardi CEL. Mood and anxiety in university students during COVID-19 isolation: a comparative study. *COVID.* 2025;5(8):127. doi:10.3390/covid5080127
12. Voss C, Shorter P, Weatrowski G, Mueller-Coyne J, Turner K. A comparison of anxiety levels before and during the COVID-19 pandemic. *Psychol Rep.* 2022;126(6):2669-2689. doi:10.1177/00332941221093250
13. Fouda SM, Madi M, Ellakany P, et al. Perception of dental students and educators about e-learning during COVID-19 pandemic. *BMC Oral Health.* 2025;25(1):829. doi:10.1186/s12903-025-06152-6
14. Csorba LM, Dabija DC. Impact of the COVID-19 pandemic on students' future online education behaviour. *Heliyon.* 2024;10(20):e39560. doi:10.1016/j.heliyon.2024.e39560
15. Özgüç S, Kaplan Serin E, Tanriverdi D. Death anxiety associated with COVID-19: a systematic review and meta-analysis. *Omega.* 2021;88(3):823-856. doi:10.1177/00302228211050503
16. Fioravanti-Bastos ACM, Cheniaux E, Landeira-Fernandez J. Development and validation of a short-form version of the Brazilian State-Trait Anxiety Inventory. *Psicol Reflex Crit.* 2011;24(3):485-494. doi:10.1590/s0102-79722011000300009
17. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Lancet.* 2007;370(9596):1453-1457. doi:10.1016/S0140-6736(07)61602-X
18. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res.* 2004;6(3):e34. doi:10.2196/jmir.6.3.e34
19. Associação Brasileira de Empresas de Pesquisa (ABEP). Critério de Classificação Econômica Brasil 2019. São Paulo: ABEP; 2019. Available from: https://www.abep.org/criterio-brasil
20. Biau DJ, Kernéis S, Porcher R. Statistics in brief: the importance of sample size in the planning and interpretation of medical research. *Clin Orthop Relat Res.* 2008;466(9):2282-2288. doi:10.1007/s11999-008-0346-9
21. Pucinelli CM, Oliveira AA, Nelson Filho P, et al. Online teaching, quality of life and anxiety of Brazilian dental professors during the COVID-19 outbreak. *Braz Oral Res.* 2022;36:0036. doi:10.1590/1807-3107bor-2022.vol36.0036
22. Hu C, Li F, Wang S, Gao Z, Pan S, Qing M. Role of Artificial Intelligence in personalized learning and clinical training in dental education. *Cogent Educ.* 2025;12:1. doi:10.1080/2331186x.2025.2490425
23. Junqueira E. Literacies and generative AI in education: a systematic review. *Ubiquity Proc.* 2025;6(1):15. doi:10.5334/uproc.183
24. Agius A, Gatt G, Zahra EV, et al. Self-reported dental student stressors and experiences during the COVID-19 pandemic. *J Dent Educ.* 2021;85(2):208-215. doi:10.1002/jdd.12409
25. Rodríguez-Molinero J, Corral-Liria I, Jiménez-Fernández R, et al. Resilience and psychological factors among dentistry students receiving face-to-face lectures during the COVID-19 pandemic. *BMC Med Educ.* 2024;24(1):446. doi:10.1186/s12909-024-05445-8
26. Asmoro PS. Psychosocial learning environment, learning anxiety, and learning satisfaction among distance learners during COVID-19. *Cogent Educ.* 2024;11(1):2359870. doi:10.1080/2331186x.2024.2359870
27. Moradi S, Fateh MS, Movahed E, et al. Prevalence of depression, anxiety, and sleep disorder among dental students: a systematic review and meta-analysis. *J Dent Educ.* 2024;88(7):900-909. doi:10.1002/jdd.13506
28. Sahin S, Tuna R. Effect of anxiety on thriving levels of university students during the COVID-19 pandemic. *Collegian.* 2022;29(3):263-270. doi:10.1016/j.colegn.2021.10.004

Journal of Dental Sciences and Education

Original Article

DOI: 10.51271/JDSE-0085

Digital literacy and computer self-efficacy among dental students and faculty members

Resul İlhan[1], Oğuz Çetin*[2]

[1]Department of Mathematics and Science Education, Institute of Educational Sciences, Niğde Ömer Halisdemir University, Niğde, Turkiye

[2]Department of Mathematics and Science Education, Faculty of Education, Niğde Ömer Halisdemir University, Niğde, Turkiye

Cite this article as: İlhan R, Çetin O. Digital literacy and computer self-efficacy among dental students and faculty members. *J Dent Sci Educ.* 2026;4(3):105-118. doi:10.51271/JDSE-0085

Received: 13.05.2026 ◆ **Accepted**: 17.08.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Aims: This study aimed to examine the digital literacy and computer self-efficacy of faculty members and undergraduate students in dental education and to investigate their associations with demographic and technology-related characteristics.

Methods: This cross-sectional survey included 102 undergraduate dental students and all 12 faculty members employed at the institution during the study period. Data were collected using the Personal Information Form, Questionnaire for Determining Prospective Teachers' Digital Literacy Levels, and Computer Self-Efficacy Scale. Descriptive statistics, independent-samples t-tests, Welch's t-test, one-way analysis of variance, non-parametric tests, correlation analyses, and multiple linear regression analyses were performed as appropriate.

Results: Male students had higher digital literacy and computer self-efficacy scores than female students, whereas no significant differences were observed between age groups. Personal computer ownership and previous computer-related training were associated with higher scores on both outcomes. In the multivariable analyses, gender, personal computer ownership, and previous computer-related training remained significantly associated with both digital literacy and computer self-efficacy after adjustment for age, grade level, and frequency of internet use. Grade level was not independently associated with either outcome after adjustment. Faculty members had higher digital literacy and computer self-efficacy scores than students. Digital literacy and computer self-efficacy were strongly positively correlated among students, whereas no statistically significant association was identified among faculty members.

Conclusion: Digital literacy and computer self-efficacy among dental students were associated with gender, access to a personal computer, and previous computer-related training. These findings support the integration of structured technology-supported learning opportunities into dental education while emphasizing that the observed associations should not be interpreted causally.

Keywords: Digital literacy, computer self-efficacy, dental education, higher education, educational technology, digital competence

INTRODUCTION

Rapid developments in digital technologies have transformed individuals' processes of accessing information, communication, learning and professional practice. In particular, the widespread integration of digital tools into educational environments has restructured teaching and learning processes, making online platforms, mobile technologies, simulation applications and interactive digital content central components of contemporary education.[1,2] This transformation has increased the importance of not only using technology effectively, but also accessing, evaluating, analysing and producing information ethically within digital environments.[3] Accordingly, digital literacy and computer self-efficacy have become critical competencies in modern education systems.[4]

Digital literacy is a multidimensional concept referring to individuals' ability to use digital technologies effectively, critically, safely and consciously. Beyond technical skills, digital literacy includes competencies such as information access, information evaluation, digital communication, content creation and problem-solving.[3] Today, digital literacy is considered an essential competency for academic achievement and professional success.[5] Particularly within higher education, the increasing prevalence of digital learning environments has made the digital competence levels of students and faculty members increasingly important.[6]

Computer self-efficacy, on the other hand, refers to individuals' perceptions of their competence in using computers and digital technologies. Based on Bandura's self-efficacy theory, computer self-efficacy is regarded as one of the key variables influencing individuals' motivation regarding technology use, their attitudes toward digital environments and their confidence in performing technology-related tasks.[7] The literature suggests that individuals with stronger self-efficacy beliefs demonstrate more positive attitudes toward technology

***Corresponding Author:** Oğuz Çetin, oguzcetin@ohu.edu.tr

use, adapt more easily to digital learning environments and perform more successfully in technology-supported applications.[8,9]

Previous studies have also demonstrated a significant relationship between digital literacy and computer self-efficacy. Individuals who can use digital tools effectively tend to develop stronger confidence regarding technology use, whereas individuals with high self-efficacy beliefs are more likely to engage actively with digital environments and adapt more successfully to technology-supported learning processes.[4,9,10] Therefore, digital literacy and computer self-efficacy may be considered mutually reinforcing constructs that contribute to individuals' successful participation in contemporary educational environments.

The increasing integration of digital technologies into health education has further emphasized the importance of digital competence among both students and faculty members. Dental education, in particular, represents one of the health disciplines in which digital technologies are used extensively. Digital radiography systems, CAD/CAM applications, virtual simulations, intraoral scanners, three-dimensional modelling technologies and online learning platforms have become integral components of contemporary dental education.[11,12] These technologies not only facilitate the development of students' theoretical and clinical skills but also require faculty members to develop advanced digital pedagogical competencies.[13]

Recent research has further demonstrated the growing importance of digital competence within dental education. Studies involving dental students have examined their experiences with CAD/CAM systems, intraoral scanners, digital workflows, and other technology-supported learning environments, generally indicating positive attitudes toward digital dentistry while also identifying variability in practical experience, confidence, and preparedness for technology-supported clinical practice.[14,16] Recent evidence also suggests that structured exposure to digital technologies during dental education may be relevant to students' confidence and preparedness for digital clinical workflows.[15,16] At a broader level, recent systematic reviews demonstrate that digitalization is becoming increasingly integrated into undergraduate and postgraduate dental education, although implementation and educational outcomes remain heterogeneous across institutions and technologies.[17,18]

More recent research has begun to examine digital competence itself among dental students rather than focusing exclusively on individual dental technologies. Shooriabi et al.[19] assessed dental students using a DigComp 2.2-based instrument and reported relatively high self-perceived digital competence, while overall competence did not differ significantly according to gender, age, academic year, or duration of daily internet use. These findings illustrate the growing interest in broader digital competence within dental education while also suggesting that its demographic correlates may vary across educational contexts.

Technology-supported educational environments, including online learning systems, virtual reality applications and simulation-based instructional tools, contribute to the development of students' clinical competencies and support more interactive and individualized learning experiences.[20] However, the effective integration of digital technologies into educational processes is closely associated with individuals' levels of digital literacy and their confidence regarding computer use.[21] Consequently, determining the digital competence levels of students and faculty members has become increasingly important for improving the quality of dental education and supporting effective technology integration within professional training environments.[22]

Although the literature on digitalization in dental education has expanded considerably in recent years, much of the existing research has focused on students' perceptions, experiences, or performance regarding specific technologies such as CAD/CAM systems, intraoral scanners, virtual simulation, and other digital workflows.[14,18] Recent research has begun to examine broader digital competence and digital literacy among dental students,[19,23] but comparatively limited attention has been given to digital literacy and computer self-efficacy as related but distinct constructs within dental education, particularly through the simultaneous inclusion of students and faculty members.[22] Although previous dental education research has compared students and faculty members in related areas such as e-learning readiness,[30] digital literacy and computer self-efficacy have received substantially less attention as distinct but potentially interrelated constructs examined concurrently across these two groups. Moreover, evidence regarding the demographic and technology-related characteristics independently associated with these constructs among dental students remains limited.[19,23]

Accordingly, the present study aimed to examine digital literacy and computer self-efficacy among dental students and faculty members within the same institutional context and to investigate the relationship between these constructs. In addition, the study examined their associations with demographic and technology-related characteristics among students and used multivariable analyses to determine which factors remained independently associated with digital literacy and computer self-efficacy after adjustment for other measured characteristics. By examining both constructs and both participant groups within a single dental education setting, the study seeks to contribute to a more comprehensive understanding of digital competence and technology-related confidence in contemporary dental education.

METHODS

Ethics

This study was approved by the Niğde Ömer Halisdemir University Ethics Committee (Date: 29.03.2024, Decision No: 2024/07-16). All procedures were conducted in accordance with the Declaration of Helsinki.

Study Design

This study employed a cross-sectional survey design[24] to examine the digital literacy levels and computer self-efficacy beliefs of dental students and faculty members. Within this framework, the relationships between digital literacy, computer self-efficacy and various demographic variables were analysed.

Participants

The target population of the study consisted of all undergraduate students and faculty members at the Faculty of Dentistry, Niğde Ömer Halisdemir University, during the data collection period. A total population approach was adopted, and all eligible students and faculty members were invited to participate in the study rather than selecting a subgroup through a sampling procedure.[25]

At the time of data collection, 121 undergraduate students were enrolled in the faculty. Of these students, 102 provided complete and valid responses and were included in the final analysis, corresponding to a valid response rate of 84.3%. The faculty employed 12 academic staff members during the same period, all of whom participated in the study, corresponding to a response rate of 100%. Thus, the final analytical sample consisted of 114 participants (102 students and 12 faculty members).

At the time of the study, the faculty of dentistry was newly established and provided education only to first-, second-, and third-year students. Fourth- and fifth-year cohorts had not yet been established at the institution. Therefore, fourth- and fifth-year students were not excluded on the basis of any eligibility or sampling criterion; rather, these cohorts did not exist during the study period. This institutional characteristic should be considered when interpreting comparisons across year levels.

An a priori sample size or power calculation was not performed because the study was designed to invite the entire eligible population of the faculty during the data collection period rather than to select a predetermined sample. Nevertheless, the small absolute number of faculty members (n=12) should be considered when interpreting analyses involving this group.

Of the students, 64 were female (62.7%) and 38 were male (37.3%); the overall distribution of the study group, including the faculty members, was evaluated within the scope of the research.

Data Collection Tools

The Personal Information Form, Questionnaire for Determining Prospective Teachers' Digital Literacy Levels and the Computer Self-Efficacy Scale were used as data collection tools in the study.

Personal Information Form: The personal information form, developed by the researchers, consists of questions designed to determine demographic variables such as gender, age, year group, ownership of a personal computer, frequency of internet use, and whether the respondent has received computer-related training.

Questionnaire for Determining Prospective Teachers' Digital Literacy Levels: Digital literacy was assessed using the instrument developed by Kıyıcı[26] to determine prospective teachers' digital literacy levels. The instrument was developed following a review of the relevant literature and existing questionnaires. An initial pool of 128 Likert-type items addressing digital literacy and its dimensions was prepared. The draft instrument was reviewed by eight field experts in terms of content appropriateness, clarity of the instructions and items, and response format. Based on expert feedback, 16 items were removed, and the revised instrument was subsequently pilot-tested with 251 prospective teachers. The internal consistency coefficient obtained in the pilot study was Cronbach's α=.9798. Following item analysis, items with unsuitable corrected item-total correlations were removed, resulting in a final 107-item instrument. The instrument addresses digital literacy through four dimensions: computer literacy, information literacy, technology literacy, and media literacy. Items are rated on a five-point frequency-based response format ranging from never to always, with intermediate categories of rarely, sometimes, and often. Scores are calculated using the arithmetic mean of the relevant items, with higher scores indicating higher levels of digital literacy. In the present study, the Cronbach's α coefficient was calculated as 0.984, indicating a high level of internal consistency.

Computer Self-Efficacy Scale: The Computer Self-Efficacy Scale developed by Şensoy[27] was used to assess participants' perceived self-efficacy regarding computer use. During the development of the instrument, an initial pool of 20 items was prepared. Following expert review for content relevance, the instrument was revised to 16 items. The 16-item form was subsequently pilot-tested with 109 teachers. Item analyses supported the retention of all 16 items, and the scale demonstrated high internal consistency, with a Cronbach's α coefficient of .89. The instrument uses a five-point Likert-type response format ranging from 1 (strongly disagree) to 5 (strongly agree), with higher scores indicating higher perceived computer self-efficacy. In the present study, the Cronbach's α coefficient was calculated as 0.814, indicating acceptable reliability.

Data Collection

The data collection phase of the study was carried out during the 2024-2025 academic year. The data collection tools were administered to faculty members and students who volunteered to participate in the study using both paper-based and electronic formats. Participants were informed about the purpose of the study, and informed consent was obtained prior to data collection. Completion of the questionnaires took approximately 15 minutes. The data obtained from participants were handled in accordance with the principle of confidentiality. Before the statistical analyses were conducted, the questionnaires were checked for completeness, and only complete responses were included in the analysis.

Statistical Analysis

Prior to inferential analyses, the distributions of the study variables were examined using the Shapiro-Wilk test, skewness and kurtosis values, and inspection for potential outliers. Homogeneity of variance for group comparisons was assessed using Levene's test. For two-group comparisons, student's independent-samples t-test was used when the homogeneity-of-variance assumption was satisfied, whereas Welch's t-test was used when this assumption was violated. For comparisons involving more than two groups, one-way ANOVA was used when the relevant assumptions were satisfied. When the omnibus one-way ANOVA was statistically significant, Tukey's honestly significant difference (HSD) test was used for post-hoc pairwise comparisons, with multiplicity-adjusted p-values reported.

For the student sample, both digital literacy and computer self-efficacy scores showed acceptable distributional characteristics; therefore, Pearson's correlation coefficient was used to examine their association. In contrast, given the small faculty sample (n=12), the significant departure from normality observed for computer self-efficacy scores, and the presence of a potential outlying observation, Spearman's rank-order correlation coefficient was used to examine the association between digital literacy and computer self-efficacy among faculty members. Statistical significance was set at $p<.05$.

In the results of the independent samples t-test, Cohen's d coefficient was calculated to determine the effect size. In the interpretation of Cohen's d-values, 0.20 was considered a small, 0.50 a medium, and 0.80 a large effect size.[28] In one-way analyses of variance, the eta squared (η^2) coefficient was used to determine the effect size. In the interpretation of η^2 values, 0.01 was considered a small effect size, 0.06 a medium effect size, and 0.14 a large effect size.[28]

To account for potential confounding among demographic and technology-related characteristics, two multiple linear regression models were additionally conducted for the student sample, with digital literacy and computer self-efficacy entered separately as dependent variables. Gender, age, grade level, personal computer ownership, frequency of internet use, and previous computer-related training were entered simultaneously as predictors.

Given that multiple hypothesis tests were conducted across several demographic and technology-related variables, the possibility of an increased family-wise type I error rate was considered. These subgroup analyses were treated as exploratory rather than confirmatory; therefore, no formal multiplicity adjustment was applied. Statistical significance was evaluated at the conventional α=.05 level, but findings with p-values close to this threshold were interpreted cautiously and in conjunction with their corresponding effect sizes. Accordingly, these results should be regarded as hypothesis-generating and warrant confirmation in future studies.

RESULTS

Demographic Findings for the Sample Group

A total of 114 participants took part in the study, including 102 students and 12 faculty members. Among the students, 62.7% were female (n=64) and 37.3% were male (n=38). Most participants were between the ages of 17 and 19. In terms of grade level, 43.1% of the students were in year 1, 35.3% were in year 2 and 21.6% were in year 3 (**Table 1**).

Table 1. Demographic characteristics of students in the sample group

Characteristic		f	%
Gender	Female	64	62.7
	Male	38	37.3
Age	17-19	54	52.9
	20 +	48	47.1
Class	Year 1	44	43.1
	Year 2	36	35.3
	Year 3	22	21.6
Ownership of a personal computer	Yes	58	56.9
	No	44	43.1
Frequency of internet use	Medium	14	13.7
	Frequent	50	49.0
	Very frequent	38	37.3
Status of computer-related training	Yes	33	32.4
	No	69	67.6
Students' purpose of internet use	Homework preparation/ information gathering	23	22.6
	Playing games	8	7.8
	Social	65	63.8
	E-mail	3	2.9
	For professional use	3	2.9
Total		102	100

More than half of the students owned a personal computer (56.9%), while 49% reported frequent internet use. However, 67.6% stated that they had not received any computer-related training previously. Regarding the purpose of internet use, most students reported using the internet mainly for social purposes.

Table 2 shows that faculty members primarily used the internet for information gathering and e-mail communication. Both groups reported regular engagement with digital technologies. Nevertheless, the proportion of students who had received computer-related training remained relatively low.

Table 2. Purposes of internet use among faculty members in the sample group

Characteristic		f	%
Purpose of internet use by faculty members	Information gathering	6	50
	Playing games	0	0
	For social purposes	2	16.6
	E-mail	4	33.3
	For professional use	0	0
Total		12	100

Findings on Normality

As shown in **Table 3**, neither digital literacy (p=.061) nor computer self-efficacy (p=.373) scores showed a statistically significant departure from normality among students. Skewness and kurtosis values were also within acceptable limits.

Table 3. Results of the normality assessment for digital literacy and computer self-efficacy scores among students

						Shapiro-Wilk	
Variable	X̄	SD	Median	Skewness	Kurtosis	Statistics	p
Digital literacy	3.46	.72	3.31	.238	-.594	.976	.061
Computer self-efficacy	3.51	.58	3.53	-.259	-.352	.986	.373

p>.05 indicates no statistically significant departure from normality. SD: Standard deviation

As shown in **Table 4**, digital literacy scores did not significantly deviate from normality (p=.894), whereas computer self-efficacy scores showed a statistically significant departure from normality (p=.042). Given the small faculty sample (n=12), this distributional characteristic was taken into account when selecting the subsequent statistical procedures.

Table 4. Results of the normality assessment for digital literacy and computer self-efficacy scores among faculty members

						Shapiro-Wilk	
Variable	X̄	SD	Median	Skewness	Kurtosis	Statistics	p
Digital literacy	3.93	.56	3.91	-.097	-.641	.968	.894
Computer self-efficacy	3.91	.48	4.03	-1.114	.898	.855	.042*

p<.05 indicates a statistically significant departure from normality. *Significant at the p<.05 level. SD: Standard deviation

Findings Regarding Digital Literacy and Computer Self-Efficacy Levels

Table 5 shows that the mean digital literacy score among students was 3.46 (SD=0.72), while the mean computer self-efficacy score was 3.51 (SD=0.58). Because validated cutoff values for categorizing these scale scores as low, moderate, or high are not available, the scores are presented and interpreted as continuous measures rather than categorical levels.

Table 5. Distribution of students' scores on digital literacy levels and attitudes towards computer self-efficacy beliefs

Variable	n	X̄	Mode	Median	SD	Observed range	Possible range
Digital literacy	102	3.46	2.76	3.310	.72	1.82-5.00	1.00-5.00
Computer self-efficacy	102	3.51	3.50	3.531	.58	1.94-4.62	1.00-5.00

SD: Standard deviation

Table 6 indicates that faculty members had higher mean scores for both digital literacy and computer self-efficacy than students. In addition, faculty members' scores showed a narrower observed range.

Table 6. Distribution of scores for faculty members' digital literacy levels and attitudes towards computer self-efficacy beliefs

Variable	n	X̄	Mode	Median	SD	Observed range	Possible range
Digital literacy	12	3.93	3.06	3.915	.56	3.06-4.86	1.00-5.00
Computer self-efficacy	12	3.91	4.25	4.031	.48	2.81-4.37	1.00-5.00

SD: Standard deviation

Findings Regarding the Gender Variable

Table 7 shows a statistically significant gender difference in students' digital literacy scores, Welch's t (59.92)=-2.220, p=.030, with male students reporting higher scores than female students. A statistically significant gender difference was also observed for computer self-efficacy, t (100)=-2.432, p=.017, again in favour of male students.

Table 7. Comparison of students' digital literacy and computer self-efficacy scores according to gender

Variable	Gender	n	X̄	SD	df	Test statistic	p	Cohen's d
Digital literacy	Female	64	3.33	.61	59.92	-2.220[a]	.030*	.47
	Male	38	3.68	.84				
Computer self-efficacy	Female	64	3.41	.52	100	-2.432[b]	.017*	.50
	Male	38	3.69	.64				

[a]Welch's t-test was used because the homogeneity-of-variance assumption was violated. [b]Student's independent-samples t-test was used because the homogeneity-of-variance assumption was satisfied. *Significant at the p<.05 level. SD: Standard deviation

Findings Regarding the Age Variable

As shown in **Table 8**, students aged 17-19 years did not differ significantly in digital literacy from those aged 20 years and over, t (100)=0.772, p=.442, Cohen's d=0.15. Similarly, no statistically significant difference was observed in computer self-efficacy between the two age groups, t (100)=-0.567, p=.572, Cohen's d=-0.11. The small effect sizes further indicate that age-related differences in both outcomes were limited.

Table 8. Comparison of students' digital literacy and computer self-efficacy scores according to age

Variable	Age	n	X̄	SD	df	t	p	Cohen's d
Digital literacy	17-19	54	3.51	.78	100	0.772	.442	.15
	20+	48	3.40	.65				
Computer self-efficacy	17-19	54	3.48	.62	100	-0.567	.572	-.11
	20+	48	3.55	.54				

SD: Standard deviation

Findings Regarding Class Level

Table 9 shows that students' digital literacy scores did not differ significantly across grade levels, F (2, 99)=2.097, p=.128, η^2=.041. In contrast, computer self-efficacy scores differed significantly according to grade level, F (2, 99)=3.263, p=.042, η^2=.062. Third-year students had the highest mean computer self-efficacy score (X̄ =3.73, SD=0.51), followed by first-year (X̄ =3.54, SD=0.58) and second-year students (X̄ =3.35, SD=0.59). Tukey HSD post-hoc comparisons indicated that the difference between second- and third-year students was statistically significant (mean difference=0.387, adjusted p=.035), whereas the differences between first- and second-year students (adjusted p=.281) and between first- and third-year students (adjusted p=.403) were not statistically significant. Therefore, although third-year students had the highest mean computer self-efficacy score, the findings do not indicate a gradual increase in computer self-efficacy with academic progression.

Table 9. Comparison of students' digital literacy and computer self-efficacy scores according to grade level

Variable	Class level	n	X̄	SD	f	p	η^2
Digital literacy	Year 1	44	3.56	.73	2.097	.128	.041
	Year 2	36	3.27	.70			
	Year 3	22	3.59	.70			
Computer self-efficacy	Year 1	44	3.54	.58	3.263	.042*	.062
	Year 2	36	3.35	.59			
	Year 3	22	3.73	.51			

Tukey HSD post-hoc comparisons were conducted following the significant omnibus ANOVA for computer self-efficacy. Only the year 2-year 3 comparison was statistically significant (adjusted p=.035); year 1-year 2 (adjusted p=.281) and year 1-year 3 (adjusted p=.403) were not significant. *η^2=eta squared. Significant at the p<.05 level. SD: Standard deviation, HSD: Honestly significant difference, ANOVA: Analysis of variance

Findings Regarding Ownership of a Personal Computer

Table 10 shows that students who owned a personal computer had significantly higher digital literacy scores (X̄ =3.64, SD=0.69) than those who did not own a personal computer (X̄ =3.23, SD=0.70), t (100)=2.922, p=.004, Cohen's d=0.58. Similarly, students who owned a personal computer reported significantly higher computer self-efficacy scores (X̄ =3.67, SD=0.50) than those without a personal computer (X̄ =3.31, SD=0.62), t (100)=3.201, p=.002, Cohen's d=0.64. These findings indicate that personal computer ownership was associated with higher scores on both digital literacy and computer self-efficacy.

Table 10. Comparison of students' digital literacy and computer self-efficacy scores according to personal computer ownership

Variable	Computer ownership	n	X̄	SD	df	t	p	Cohen's d
Digital literacy	Yes	58	3.64	.69	100	2.922	.004*	.58
	No	44	3.23	.70				
Computer self-efficacy	Yes	58	3.67	.50	100	3.201	.002*	.64
	No	44	3.31	.62				

Cohen's d was calculated using the pooled standard deviation. *Significant at the p<.05 level. SD: Standard deviation

Findings Regarding Frequency of Internet Use

The Kruskal-Wallis test examined whether students' digital literacy and computer self-efficacy scores differed according to the frequency of internet use. **Table 11** shows that the rank averages for digital literacy scores were similar across students who used the internet moderately, frequently and very frequently. The analysis therefore revealed no significant difference between the groups (X^2=0.167, p=.920).

Table 11. Comparison of students' digital literacy and computer self-efficacy scores according to internet usage frequency

Variable	Frequency	n	Rank mean	df	X^2	p
Digital literacy	Medium	14	49.14	2	.167	.920
	Frequent	50	51.18			
	Very frequent	38	52.79			
Computer self-efficacy	Medium	14	46.79	2	.591	.744
	Frequent	50	53.41			
	Very frequent	38	50.72			

Similarly, computer self-efficacy scores did not differ significantly based on internet usage frequency (X^2=0.591, p=.744). Although students who used the internet frequently reported slightly higher rank averages, the differences between the groups remained statistically insignificant. Overall, internet usage frequency did not produce a significant difference in students' digital literacy or computer self-efficacy levels.

Findings Regarding the Status of Computer-Related Training

Table 12 shows that students who had received previous computer-related training reported significantly higher digital literacy scores ($\bar{X}$=3.71, SD=0.61) than those who had not received such training ($\bar{X}$=3.34, SD=0.74), t (100)=2.455, p=.016, Cohen's d=0.52. Similarly, students with previous computer-related training demonstrated significantly higher computer self-efficacy scores ($\bar{X}$=3.80, SD=0.49) than students without such training ($\bar{X}$=3.38, SD=0.58), t (100)=3.637, p<.001, Cohen's d=0.77. These findings indicate that previous computer-related training was associated with higher levels of both digital literacy and computer self-efficacy. The effect size was moderate for digital literacy (Cohen's d=0.52) and moderate-to-large for computer self-efficacy (Cohen's d=0.77).

Table 12. Comparison of students' digital literacy and computer self-efficacy scores according to previous computer-related training

Variable	Receipt of training	n	$\bar{X}$	SD	df	t	p	Cohen's d
Digital literacy	Yes	33	3.71	.61	100	2.455	.016*	.52
	No	69	3.34	.74				
Computer self-efficacy	Yes	33	3.80	.49	100	3.637	<.001*	.77
	No	69	3.38	.58				

Cohen's d was calculated using the pooled standard deviation, *Significant at the p<.05 level. SD: Standard deviation

Findings Regarding the Comparison of Digital Literacy and Computer Self-Efficacy Levels Among Faculty Members and Students

Table 13 presents the results of the Mann-Whitney U test comparing the digital literacy levels of students and faculty members. Faculty members obtained higher rank averages than students, and the difference between the groups was statistically significant (U=372.500, p=.027, r=.207). These findings suggest that faculty members demonstrated higher digital literacy levels than students.

Table 13. Comparison of digital literacy scores of students and faculty members

Variable	Status	n	Rank mean	Total	U	p	Effect size r
Digital literacy	Student	102	55.15	5625.50	372.500	.027*	.207
	Faculty members	12	77.46	929.50			

Effect size r was calculated as |Z|/√N. *Significant at the p<.05 level

Table 14 shows a similar pattern for computer self-efficacy scores. Faculty members again reported higher rank averages than students, and the difference between the groups reached statistical significance (U=362.000, p=.021, r=.216). Accordingly, faculty members demonstrated significantly higher computer self-efficacy levels compared to students.

Table 14. Comparison of scores for students' and faculty members' computer self-efficacy beliefs

Variable	Status	n	Rank mean	Total	U	p	Effect size r
Computer self-efficacy	Student	102	55.05	5615.00	362.000	.021*	.216
	Faculty members	12	78.33	940.00			

Effect size r was calculated as |Z|/√N. *Significant at the p<.05 level

Findings Regarding the Relationship Between Digital Literacy and Computer Self-Efficacy Beliefs

Given the small faculty sample and the departure from normality observed in computer self-efficacy scores, Spearman's rank-order correlation was used to examine the relationship between digital literacy and computer self-efficacy. As shown in Table 15, a weak positive but statistically non-significant association was observed between the two variables (ρ=.177, p=.583).

Table 15. Spearman's rank-order correlation between digital literacy and computer self-efficacy among faculty members

Variables	n	Spearman's ρ	p
Digital literacy-computer self-efficacy	12	.177	.583

Spearman's rank-order correlation was used because of the small faculty sample and the departure from normality observed in computer self-efficacy scores

Table 16 presents the relationship between students' digital literacy levels and computer self-efficacy beliefs. The analysis revealed a strong positive and statistically significant correlation between the two variables (r=.660, p<.001). In other words, students with higher digital literacy levels also reported stronger computer self-efficacy beliefs.

Table 16. The relationship between students' digital literacy levels and computer self-efficacy beliefs

Correlations		Digital literacy	Computer self-efficacy
Digital literacy	r	1	.660**
	p		<.001
	n	102	102
Computer self-efficacy	r	.660**	1
	p	<.001	
	n	102	102

*Significant at the p<.05 level

Multivariable Analyses

To examine whether the observed associations remained after accounting for potentially interrelated participant characteristics, two multiple linear regression analyses were conducted in the student sample. Gender, age, grade level, personal computer ownership, frequency of internet use, and previous computer-related training were entered simultaneously as predictors.

As shown in **Table 17**, the overall model predicting digital literacy was statistically significant, F (8, 93)=3.229, p=.003, explaining 21.7% of the variance in digital literacy scores (R^2=.217; adjusted R^2=.150). After simultaneous adjustment for the other variables in the model, gender (β=-.247, p=.013), personal computer ownership (β=-.246, p=.012), and previous computer-related training (β=-.221, p=.026) remained significantly associated with digital literacy. Female students had lower adjusted digital literacy scores than male students, while students without a personal computer and those without previous computer-related training also had lower adjusted scores. Age, grade level, and frequency of internet use were not independently associated with digital literacy.

Table 17. Multiple linear regression analysis predicting digital literacy among students

Predictor	B	β	p
Gender (female vs. male)	-0.365	-.247	.013*
Age (≥20 vs. 17-19 years)	-0.066	-.046	.700
Grade level (year 2 vs. year 1)	-0.314	-.210	.072
Grade level (year 3 vs. year 1)	-0.165	-.095	.483
Personal computer ownership (no vs. yes)	-0.355	-.246	.012*
Internet use (frequent vs. medium)	0.117	.081	.585
Internet use (very frequent vs. medium)	0.147	.099	.508
Previous computer-related training (no vs. yes)	-0.338	-.221	.026*

Reference categories were male, 17-19 years of age, year 1, personal computer ownership, medium internet use, and previous computer-related training. R^2=.217; adjusted R^2=.150. Overall model: F (8, 93)=3.229, p=.003. *Significant at the p<.05 level

As presented in **Table 18**, the model predicting computer self-efficacy was also statistically significant, F (8, 93)=4.757, p<.001, accounting for 29.0% of the variance (R^2=.290; adjusted R^2=.229). Gender (β=-.229, p=.015), personal computer ownership (β=-.288, p=.002), and previous computer-related training (β=-.290, p=.003) remained significantly associated with computer self-efficacy after simultaneous adjustment. Female students, students without a personal computer, and students without previous computer-related training had lower adjusted computer self-efficacy scores. Age, grade level, and frequency of internet use were not independently associated with computer self-efficacy.

Table 18. Multiple linear regression analysis predicting computer self-efficacy among students

Predictor	B	β	p
Gender (female vs. male)	-0.274	-.229	.015*
Age (≥20 vs. 17-19 years)	0.075	.065	.568
Grade level (year 2 vs. year 1)	-0.260	-.215	.053
Grade level (year 3 vs. year 1)	-0.087	-.062	.629
Personal computer ownership (no vs. yes)	-0.337	-.288	.002*
Internet use (frequent vs. medium)	0.221	.191	.180
Internet use (very frequent vs. medium)	0.155	.129	.365
Previous computer-related training (no vs. yes)	-0.358	-.290	.003*

Reference categories were male, 17-19 years of age, year 1, personal computer ownership, medium internet use, and previous computer-related training. R^2=.290; adjusted R^2=.229. Overall model: F (8, 93)=4.757, p<.001. *Significant at the p<.05 level

DISCUSSION

Digital technologies have become increasingly central to both theoretical instruction and clinical training within dental education. In this context, the present study examined the digital literacy and computer self-efficacy levels of dental students and faculty members. The results showed measurable variation in digital literacy and computer self-efficacy scores across students and faculty members, with faculty members obtaining higher scores than students on both measures. These findings should be interpreted on the basis of the observed continuous scores rather than categorical levels, as validated cutoff values for classifying these measures as low, moderate, or high are not available. Similar observations have been reported in previous studies emphasizing the expanding role of online learning systems, digital simulations and technology-enhanced educational applications in health education settings.[4,20] In dentistry specifically, technologies such as digital radiography systems, CAD/CAM applications and intraoral scanners are no longer peripheral tools but increasingly routine components of professional training and clinical practice.[11]

The digital literacy and computer self-efficacy scores observed among students may be associated with the extent to which digital technologies are embedded within contemporary university life. Students today routinely engage with online learning platforms, mobile applications and digital educational resources throughout their academic experiences. Continuous interaction with such technologies may be associated with greater technological competence and confidence regarding digital tool use. Previous studies have similarly reported positive associations between sustained exposure to digital learning environments and digital competence and self-efficacy perceptions.[4,10] The present findings appear particularly relevant in dental education, where students increasingly encounter technology-intensive systems during both pre-clinical and clinical training processes.[20,29]

Recent dental education research provides additional context for these findings. Ceylan et al.[23] reported a significant positive relationship between digital literacy and perceived professional competence among dental students, highlighting the relevance of digital literacy to students' perceptions of their professional capabilities. More recently, Shooriabi et al.,[19] using a DigComp 2.2-based assessment, found that dental students demonstrated a heterogeneous profile of self-perceived digital competence, with relatively stronger perceptions in some technical and functional domains and lower competence in higher-order domains such as critical thinking, digital citizenship, and lifelong learning. Taken together, these findings suggest that digital competence among dental students should be considered a multidimensional construct rather than being inferred solely from students' routine exposure to digital technologies.

The study further revealed that faculty members demonstrated higher levels of digital literacy and computer self-efficacy than students. A similar pattern has been reported in recent dental education research. Zahid and Agou[30] compared dental students and faculty members and found that faculty members demonstrated higher scores than students across almost all domains of e-learning readiness, including computer skills, online skills, and overall readiness. Although e-learning readiness is conceptually distinct from digital literacy and computer self-efficacy, these findings provide dental education-specific evidence that students and faculty members may differ in their preparedness for technology-supported educational environments. One possible interpretation of the higher digital literacy and computer self-efficacy scores observed among faculty members may be considered within the Technological Pedagogical Content Knowledge (TPACK) framework.[31] This framework emphasizes the integration of technological, pedagogical, and content knowledge in educational practice. Faculty members may have greater opportunities than students to engage with technology in teaching, research, and professional activities. However, TPACK was not directly measured in the present study; therefore, this framework is used only as a possible theoretical perspective for interpreting the observed group differences rather than as an explanation directly supported by the study data. Faculty members may have opportunities to engage with digital tools through teaching, research, academic communication, and other professional activities; however, the frequency and nature of such technology use were not directly assessed in the present study.[32,33] From a theoretical perspective, their higher scores may be consistent with greater opportunities to engage with technological and pedagogical practices; however, the present study did not directly assess faculty members' ability to integrate pedagogical strategies with technological applications. From the perspective of Bandura's self-efficacy theory,[34] repeated successful experiences in orchestrating technology-enhanced learning environments serve as mastery experiences, thereby reinforcing their beliefs regarding their capability to perform technology-related tasks effectively. Accordingly, potential explanations involving professional experience, age, technology exposure, or pedagogical technology use should be regarded as hypotheses for future research rather than explanatory mechanisms established by the present study.

In addition, the concept of digital pedagogical competence may help explain the higher digital competence levels observed among faculty members. Digital pedagogical competence involves not only technical technology use but also the ability to integrate digital tools into teaching and learning processes effectively. Particularly after the rapid expansion of online and hybrid education models, faculty members have been required to adapt to technology-supported instructional approaches more intensively. Similar findings have been reported in the literature indicating that greater professional experience with digital technologies is positively associated with digital literacy and self-efficacy beliefs.[6,7] Nevertheless, the relatively high digital competence levels observed among students also suggest that contemporary students are increasingly familiar with digital technologies prior to entering higher education.

One of the most notable findings of the present study is the significant positive relationship identified between digital literacy and computer self-efficacy among students. This finding supports theoretical perspectives suggesting that digital competence and self-efficacy are mutually reinforcing constructs.[35] According to self-efficacy theory, individuals who successfully perform technology-related tasks develop stronger beliefs regarding their ability to use technology effectively.[34] Similarly, individuals with high levels of self-efficacy tend to engage more actively with digital technologies, persist when encountering technological difficulties and adapt more successfully to digital learning environments.[7,36] Accordingly, higher digital literacy may coexist with stronger perceptions of computer self-efficacy, while greater confidence regarding technology use may likewise be associated with more extensive engagement with digital tools. However, the cross-sectional nature of the present study does not allow the direction of this relationship to be determined.

This finding is also consistent with previous studies reporting positive relationships between digital literacy and computer self-efficacy.[9,10] However, the current study contributes to the literature by examining this relationship specifically within dental education. In dentistry, students are increasingly expected to use digital diagnostic systems, simulation technologies and technology-supported clinical applications throughout their education and professional practice. Therefore, digital literacy and computer self-efficacy may be important competencies associated with students' readiness to engage with contemporary clinical technologies and technology-supported professional environments.

Male students reported higher levels of digital literacy and computer self-efficacy than female students in the unadjusted comparisons. Importantly, the multivariable analyses showed that gender remained significantly associated with both digital literacy (β=-.247, p=.013) and computer self-efficacy (β=-.229, p=.015) after adjustment for age, grade level, personal computer ownership, frequency of internet use, and previous computer-related training. However, technology-related interests, patterns of technology use, and other potentially relevant psychosocial or contextual factors were not directly measured in the present study. Therefore, these explanations should be regarded as possible interpretations rather than mechanisms demonstrated by the present data. Thus, the observed gender differences could not be explained solely by differences in these measured characteristics. Previous studies have suggested that repeated exposure to technology-related experiences can strengthen perceptions of technological competence and self-efficacy.[8,9] At the same time, these findings should be interpreted cautiously. Recent research increasingly indicates that gender-based differences in digital competence may be narrowing as access to digital technologies and participation in online learning environments become more widespread.[4]

However, recent evidence from dental education does not consistently demonstrate gender-related differences in digital competence. Shooriabi et al.[19] found no statistically significant difference in overall self-perceived digital competence between male and female dental students (p=.73), with a negligible effect size (Cohen's d=.05). Similarly, a large study of undergraduate dental students examining ICT-related skills and technology use reported no significant gender differences in overall ICT skills or educational use of digital technologies.[37] These findings contrast with the present study, in which male students had higher digital literacy and computer self-efficacy scores and gender remained independently associated with both outcomes after adjustment for the other measured characteristics. This discrepancy suggests that gender-related patterns in digital competence may vary across institutional, educational, and sociocultural contexts and should not be assumed to represent a consistent characteristic of dental students.

Although gender remained independently associated with both outcomes in the adjusted analyses, this association should not be interpreted as evidence of a causal effect of gender. Other unmeasured factors, including the nature and quality of previous digital experiences, technology usage patterns, socio-cultural influences, individual interest in technology, and opportunities for engagement with digital learning environments, may contribute to the observed differences. The cross-sectional design of the present study also precludes causal inference.

Within the context of dental education, this issue becomes particularly important because contemporary dental training increasingly requires all students, regardless of gender, to engage actively with technology-intensive systems such as digital imaging applications, virtual simulations, CAD/CAM systems and computer-assisted treatment planning tools.[29,38] Consequently, educational initiatives aimed at reducing possible inequalities in digital competence development may contribute to more equitable participation in technology-supported clinical education processes.[33] From this perspective, minimizing gender-based disparities in access to and use of digital technologies may support more inclusive and sustainable professional training environments.[39]

No significant differences were identified between the 17-19-year-old and ≥20-year-old student groups in terms of digital literacy or computer self-efficacy. This result may be related to the relatively similar digital experiences shared by contemporary university students. Regardless of age, students are now routinely exposed to online learning systems, mobile technologies and digital communication platforms throughout their education. As a result, age may no longer function as a strong differentiating factor in technology-related competence within higher education environments. Similar findings have been reported in previous studies suggesting that digital technologies have become integrated into the everyday educational experiences of younger generations.[10,40] Accordingly, age-related differences in digital competence may have become less pronounced compared with previous periods.[41] In contemporary higher education settings, access to digital technologies and routine engagement with online platforms may contribute to more homogeneous levels of digital experience among students.[42]

This finding is consistent with several previous studies reporting that age does not necessarily create significant differences in digital literacy or computer self-efficacy among university students.[10] Nevertheless, some studies have argued that older individuals may experience greater difficulties in adapting to emerging technologies.[3] The absence of significant age differences in the current study may therefore be explained by the relatively homogeneous age distribution of the participants and their similar exposure to digital learning environments within higher education.

Recent dental education research provides further support for the absence of a clear age-related pattern. Shooriabi et al.[19] found no statistically significant difference in overall self-perceived digital competence among dental students according to age. This dental education-specific finding is consistent with the present results, in which neither the unadjusted comparisons nor the multivariable analyses identified a significant association between age and digital literacy or computer self-efficacy. Nevertheless, given the relatively restricted age range of the present sample, these findings should not be interpreted as evidence that age is unrelated to digital competence across broader or more age-diverse dental populations.

From the perspective of health education, this finding may indicate that digital competence is increasingly becoming a universal educational requirement rather than an age-dependent skill.[33,43] Particularly in dental education, students are expected to utilise digital systems actively regardless of age due to the increasing prevalence of technology-supported diagnostic and clinical applications.[29] Consequently, initiatives aimed at improving digital competence should be designed inclusively to address all student groups.

The multivariable findings further supported this pattern, as age was not independently associated with either digital literacy (β=-.046, p=.700) or computer self-efficacy (β=.065, p=.568) after adjustment for the other measured demographic and technology-related characteristics.

In the unadjusted analysis, students' computer self-efficacy differed significantly according to grade level, with third-year students reporting higher levels than the other grade groups. One possible theoretical perspective for interpreting this pattern is Situated Learning Theory proposed by Lave and Wenger.[44] Although situated learning was not directly measured in the present study, this framework may help contextualize the observed grade-level pattern. As students progress through dental education, they increasingly encounter dental-specific digital systems, such as CAD/CAM technologies, intraoral scanners, and digital imaging, within authentic or simulated professional contexts.[29,38] Such experiences may transform abstract technological knowledge into applied professional competence. Repeated interaction with these technologies may also provide mastery experiences that contribute to stronger computer self-efficacy beliefs, consistent with Bandura's theoretical framework.[34] Previous research has similarly suggested that greater practical experience with technology may contribute positively to computer self-efficacy.[8,9]

However, the multivariable analysis provided a more nuanced interpretation of this finding. After adjustment for gender, age, personal computer ownership, frequency of internet use, and previous computer-related training, grade level was no longer independently associated with computer self-efficacy. Thus, although greater exposure to professional technologies among upper-year students may plausibly contribute to the pattern observed in the unadjusted analysis, the findings do not support interpreting grade level itself as an independent factor. Rather, the initial grade-level difference may partly reflect differences in students' broader technology-related experiences and access.

Recent evidence from dental education also suggests that academic progression may not necessarily correspond to higher overall digital competence. Shooriabi et al.[19] reported no statistically significant differences in overall self-perceived digital competence across academic years among dental students. Although the construct assessed in that study is not identical to computer self-efficacy, its findings reinforce the need for caution when interpreting year-of-study differences as evidence of progressive development in technology-related competence. In the present study, this caution is particularly important because the unadjusted difference in computer self-efficacy was not retained after adjustment for other demographic and technology-related characteristics.

No significant grade-level difference was observed for digital literacy in the unadjusted analysis, and grade level likewise showed no independent association with digital literacy in the multivariable model. This may suggest that students enter university with a relatively established baseline of digital competence, given the extensive integration of digital technologies into their everyday and educational lives.[40,42] Taken together, these findings indicate that progression through dental education should be considered alongside students' access to technology and previous technology-related learning experiences rather than as an isolated determinant of digital competence.

Another important finding of the study was that students who owned personal computers demonstrated significantly higher levels of digital literacy and computer self-efficacy in the unadjusted comparisons. Importantly, personal computer ownership remained significantly associated with both digital literacy (β=-.246, p=.012) and computer self-efficacy (β=-.288, p=.002) after adjustment for gender, age, grade level, frequency of internet use, and previous computer-related training. This finding highlights the potential relevance of access to technology in relation to digital competence.[35] Students who own personal computers may have greater opportunities for autonomous learning, repeated practice, and continuous interaction with digital environments.[7] According to self-efficacy theory, repeated successful experiences with technology may contribute to stronger beliefs regarding one's technological capabilities.[34] Therefore, regular engagement with digital tools may support both digital literacy and perceptions of computer self-efficacy.

The literature similarly suggests that access to digital technologies contributes positively to digital competence development.[4,21] The persistence of the association with personal computer ownership after multivariable adjustment suggests that access to a personal computer may represent an important correlate of digital competence within this sample, beyond the other measured demographic and technology-related characteristics. However, because of the cross-sectional design, this association should not be interpreted causally. This issue may be particularly relevant within dental education because of the increasing reliance on digital technologies in both theoretical and clinical training. Applications such as digital radiography systems, virtual simulations, digital patient management platforms, and three-dimensional modelling software often require continued engagement with digital devices outside classroom settings. Therefore, limited access to personal digital technologies may restrict opportunities to develop digital competence and technology-related confidence.

Dental education-specific evidence also suggests that prior experience with digital technologies may be relevant to students' engagement with digital workflows. In an undergraduate digital dentistry training module involving intraoral scanning and computer-aided design, students with previous CAD experience demonstrated better performance in CAD task completion and intraoral scanning than students without such experience.[45] Although prior CAD experience is not equivalent to personal computer ownership, this finding provides additional context for the possibility that differences in access to and experience with digital technologies may be associated with students' readiness to engage with technology-intensive dental education.

Students who had received previous computer-related training also demonstrated significantly higher levels of digital literacy and computer self-efficacy in the unadjusted comparisons. These associations remained statistically significant after adjustment for the other measured characteristics, both for digital literacy (β=-.221, p=.026) and computer self-efficacy (β=-.290, p=.003). Thus, previous computer-related training emerged as an important correlate of both outcomes even when gender, age, grade level, personal computer ownership, and frequency of internet use were considered simultaneously.

Nevertheless, this association should not be interpreted as evidence that previous computer-related training caused higher digital literacy or computer self-efficacy. Because participation in such training was not experimentally assigned, students with greater pre-existing interest in technology, greater confidence in using digital tools, or more previous technology experience may have been more likely to participate in computer-related training. Therefore, the direction of the observed association cannot be established from the present cross-sectional data.

This finding highlights a potentially important association between previous technology-based educational experiences and digital competence.[41] Previous research has suggested that computer-related training may be associated with stronger problem-solving abilities in digital environments, lower technology-related anxiety, and greater confidence regarding technology use.[7,36] From the perspective of self-efficacy theory, structured learning experiences and repeated successful interactions with technology may function as mastery experiences that reinforce individuals' beliefs regarding their technological capabilities.[34]

Recent dental education research provides more specific support for the educational relevance of structured exposure to digital technologies. Nassani et al.[14] reported significant improvements in dental students perceived knowledge and skills following a CAD/CAM educational exercise, while Liu et al.[16] found improvements in theoretical knowledge and practical performance following the integration of intraoral scanning and digital analysis tools into a dental training model. Research on full digital workflows has also indicated that students with previous CAD experience may demonstrate better performance in digital tasks such as CAD procedures and intraoral scanning.[45] Although these studies examined specific digital dentistry interventions and prior digital experience rather than the broader computer-related training variable assessed in the present study, collectively they support the educational relevance of structured and repeated exposure to digital technologies within dental curricula.

Within dental education specifically, this finding has considerable importance because contemporary dental practice increasingly depends on technology-supported diagnostic and clinical systems. The growing use of CAD/CAM systems, digital radiography, virtual simulation technologies and digital patient records requires dental students to develop advanced digital competencies during their education.[29,38] Therefore, expanding practical digital competence training within dental curricula may contribute significantly to students' professional readiness and technological adaptability.[33,43]

In contrast, frequency of internet use was not significantly associated with either digital literacy or computer self-efficacy in the unadjusted analyses. This pattern was also supported by the multivariable analyses, in which frequency of internet use was not independently associated with either outcome after adjustment for gender, age, grade level, personal computer ownership, and previous computer-related training. These findings suggest that the frequency of internet use alone may be less relevant to digital competence than the nature and purpose of technology engagement. In other words, greater frequency of internet use may not necessarily translate into more advanced digital skills or stronger confidence in computer use.

Recent dental education research provides further support for this interpretation. Shooriabi et al.[19] found no statistically significant differences in overall self-perceived digital competence among dental students according to the duration of daily internet use. This finding is consistent with the present results and suggests that the amount of time spent online may not, by itself, reflect the quality or educational relevance of students' digital engagement. Indeed, frequent use of digital technologies does not necessarily imply competence across more complex domains such as information evaluation, critical thinking, digital content creation, or responsible participation in digital environments.[19]

Among faculty members, Spearman's rank-order correlation indicated a weak positive but statistically non-significant association between digital literacy and computer self-efficacy ($\rho=.177$, $p=.583$). This finding differs from the strong positive association observed among students. However, the faculty result should be interpreted with considerable caution because the faculty group was small (n=12), which substantially limits statistical power and the precision of the estimated association. In addition, the relatively homogeneous digital competence levels observed among faculty members may have restricted variability and, consequently, the ability to detect an association between the two constructs. It is also possible that faculty members' professional technology use is associated with institutional responsibilities and occupational requirements rather than with individual self-efficacy perceptions alone.[32] Therefore, the absence of a statistically significant association in this small faculty group should not be interpreted as evidence that digital literacy and computer self-efficacy are unrelated among faculty members. Future studies with larger faculty samples are needed to examine this relationship more reliably.

Overall, the findings of the present study suggest that digital literacy and computer self-efficacy are closely interconnected among dental students, consistent with previous research on the relationship between digital competence and self-efficacy.[35] The increasing integration of technology into dental training and clinical practice makes digital competence an essential component of professional preparation.[43] Therefore, educational environments that support students' active engagement with digital technologies may contribute not only to the development of digital literacy but also to stronger self-confidence regarding technology use.[7] In this regard, expanding technology-supported learning environments, simulation-based instructional practices and practical digital competence training within dental education programmes may support both academic achievement and professional readiness in increasingly digitalised healthcare environments.[33,38]

This study examined the digital literacy levels and computer self-efficacy beliefs of dentistry students and faculty members within the context of higher education. Furthermore, the study contributes to the growing body of literature examining digital competence within health and dental education.[29,33]

While age, grade level, and frequency of internet use were not independently associated with either outcome in the multivariable analyses, gender, personal computer ownership, and previous computer-related training remained significantly associated with both digital literacy and computer self-efficacy. These findings highlight the potential relevance of technology access and prior structured technology-related learning experiences to students' digital competence, while also indicating that the observed associations should not be interpreted causally. Furthermore, faculty members exhibited significantly higher levels across both variables compared with students; however, given the observational design and the small faculty sample, this group differences should be interpreted as associations rather than consequences of professional or pedagogical technology engagement.[32]

Implications for Dental Education

The findings of the present study have several important implications for dental education. First, the results suggest that digital competencies should be integrated more systematically into undergraduate dental curricula rather than being addressed through isolated technology-related courses. The increasing use of digital radiography systems, CAD/CAM applications, virtual simulations and digital patient management systems within contemporary dentistry requires students to develop not only technical knowledge but also confidence regarding technology use. Therefore, embedding technology-supported learning experiences into both theoretical and clinical training processes may contribute to stronger digital literacy and computer self-efficacy among dental students. From a theoretical perspective, Situated Learning Theory proposed by Lave and Wenger[44] may provide a useful framework for considering how authentic engagement with professional digital technologies could support digital competence development. However, situated learning processes were not directly assessed in the present study, and this interpretation should therefore be regarded as a theoretically informed implication rather than an empirical finding of the study.

Recent dental education research reinforces the need for more systematic curricular integration of digital technologies. Mwogosi et al.[46] identified substantial gaps in digital competency, infrastructure, and structured digital training within dental education and emphasized the combined importance of institutional readiness, faculty development, technological infrastructure, and student competency development. Similarly, recent evidence from clinical-year dental students indicates strong interest in digital dentistry alongside perceived insufficiency in existing curricular coverage, further supporting the need for structured and progressive integration of digital technologies throughout dental education.[47] These findings provide contemporary dental education-specific support for moving beyond isolated exposure toward more systematic and competency-oriented digital training.

The findings also highlight the importance of supporting continuous digital pedagogical development among faculty members. This implication is supported by recent dental education evidence demonstrating persistent limitations in faculty training for digital technologies. Mwogosi et al.[46] reported limited formal digital training among dental educators and identified faculty development as a central component of effective and sustainable digital integration in dental education. In this regard, Wenger's concept of Communities of Practice[48] may provide a valuable framework for fostering collaborative professional learning, knowledge sharing and sustainable technology integration within dental education environments. Because Communities of Practice was not directly examined in the present study, this framework is proposed as a potential approach for future faculty development rather than as an explanation empirically tested by the current data. Creating collaborative learning communities among faculty members may encourage the exchange of technological experiences, strengthen digital pedagogical competencies and support the effective adaptation of emerging educational technologies into teaching practices.

Another important implication concerns equitable access to digital technologies. The findings demonstrated that personal computer ownership and prior computer-related training were positively associated with both digital literacy and computer self-efficacy, and these associations remained statistically significant after adjustment for the other measured characteristics. The second-level digital divide framework[49] may provide one possible theoretical perspective for interpreting the observed associations between technology access, previous computer-related training, and digital competence. This framework emphasizes inequalities not only in access to digital technologies but also in the skills, experiences, and opportunities required to use these technologies effectively. However, the dimensions of the digital divide were not directly measured in the present study; therefore, this framework is used only to contextualize the findings rather than as an empirically tested explanation. Therefore, universities should implement supportive policies aimed at reducing inequalities in digital competence development and expanding students' opportunities to engage with technology-supported learning environments both inside and outside the classroom.

This implication is also supported by dental education-specific evidence. Bayazıt and Öztürk Bayazıt[50] reported that oral and dental health students with limited access to the internet and personal digital devices experienced particular difficulties with distance education and emphasized the importance of identifying and addressing digital inequalities among learners. More recent evidence from dental training institutions has likewise demonstrated disparities in access to digital infrastructure and digital tools across institutional settings.[46] Together, these findings reinforce the importance of considering equitable access alongside skills and training when planning the digital transformation of dental education.

Finally, future research should examine digital literacy and computer self-efficacy within larger and more diverse samples across different dental schools and educational contexts. Mixed-methods studies incorporating qualitative data may also provide a deeper understanding of how students and faculty members experience technology integration within dental education and how these experiences are associated with professional readiness in increasingly digitalized healthcare environments.

Limitations

Several limitations should be considered when interpreting the findings of this study. First, the study was conducted at a single institution, which may limit the generalizability of the findings to dental students and educational settings with different institutional, curricular, or technological characteristics. Although the study aimed to reach the entire eligible student and faculty populations at the institution, the findings remain specific to this institutional context. In addition, although all faculty members employed at the institution participated in the study, the absolute size of the faculty group was small (n=12) because the faculty of dentistry was newly established at the time of data collection. This small group size limits the statistical power of analyses involving faculty members, particularly correlation analyses; therefore, findings pertaining to this group should be regarded as exploratory and interpreted cautiously.

Second, because the institution had not yet established fourth- and fifth-year cohorts, the student population consisted exclusively of first-, second-, and third-year students. Consequently, the findings, particularly those concerning grade level, cannot be generalized to senior dental students, who may have greater exposure to clinical practice and digital dental technologies.

Third, both digital literacy and computer self-efficacy were assessed using self-report instruments. Accordingly, the responses may be subject to self-report bias, including differences in participants' perceptions of their own competencies, and may not necessarily reflect objectively demonstrated digital skills or technology-related performance. In addition, the instruments used in the present study were originally developed and validated in educational populations other than dental students and faculty members. Although internal consistency was assessed in the present sample, the applicability and measurement properties of these instruments specifically within dental education populations require further investigation. More specifically, although the Digital Literacy Questionnaire demonstrated high internal consistency in the present sample, this finding provides evidence of reliability rather than construct validity. Because this instrument was originally developed for prospective teachers and its construct validity has not been specifically established among dental students or dental faculty members, findings derived from this instrument should be interpreted with this measurement limitation in mind.

Fourth, the cross-sectional design of the study limits causal interpretation of the observed associations. Although multivariable analyses were conducted to account for potential confounding among gender, age, grade level, personal computer ownership, frequency of internet use, and previous computer-related training, residual confounding due to unmeasured factors cannot be excluded. Therefore, the adjusted associations identified in the present study should not be interpreted as evidence of causal relationships.

Finally, a relatively large number of statistical comparisons were conducted across demographic and technology-related characteristics. Because these analyses were exploratory, no formal adjustment for multiple testing was applied. Consequently, an increased risk of type I error cannot be excluded, particularly for findings with p-values close to the conventional significance threshold. These findings should therefore be interpreted cautiously and confirmed in future studies using larger samples, prospectively defined hypotheses, and confirmatory analytical approaches.

CONCLUSION

This study examined digital literacy and computer self-efficacy among dental students and faculty members and explored their relationships with demographic and technology-related characteristics. Digital literacy and computer self-efficacy were strongly and positively associated among students, whereas no statistically significant association was identified among faculty members; however, the latter finding should be interpreted cautiously because of the small faculty sample. The multivariable findings showed that gender, personal computer ownership, and previous computer-related training remained significantly associated with both digital literacy and computer self-efficacy after adjustment for the other measured characteristics. In contrast, age, grade level, and frequency of internet use were not independently associated with either outcome. These findings suggest that access to digital technologies and previous structured technology-related learning experiences may be particularly relevant when considering students' digital competence and confidence in technology use. Nevertheless, because of the cross-sectional design, these associations should not be interpreted as causal relationships. From an educational perspective, the findings support providing dental students with equitable opportunities to access digital technologies and participate in structured technology-supported learning experiences throughout their education. Future multicentre and longitudinal studies with larger student and faculty samples are needed to clarify the factors associated with digital competence and to examine how these competencies develop over the course of dental education.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study was approved by the Niğde Ömer Halisdemir University Ethics Committee (Date: 29.03.2024, Decision No: 2024/07-16).

Informed Consent

Written informed consent was obtained from all individual participants prior to their inclusion in the study. Participants were fully informed about the study's aims, procedures, potential risks and benefits, and their rights-including the right to withdraw at any time without consequence. All participants voluntarily signed a written informed consent form.

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: Rİ, OÇ; Design: OÇ; Control: OÇ; Resources: Rİ; Materials: Rİ; Data Collection and/or Processing: Rİ; Analysis and/or Interpretation: OÇ; Literature Review: Rİ, OÇ; Writing the Article: Rİ, OÇ; Critical Review: Rİ, OÇ.

REFERENCES

1. Çetin O, Güngör B. Computer self-efficacy beliefs of primary school teachers and their attitudes towards computer-aided instruction. *Ondokuz Mayıs Univ J Fac Edu.* 2014;33(1):55-77. doi:10.7822/omuefd.214866
2. Kır Ş. The evolving roles of higher education institutions and academic staff in the digital transformation process. *J Open Edu Appl Res.* 2020; 6(3):143-163.
3. Kaleci D. Digital literacy. In: Aydemir H, Ciğerci FM, Karalı Y, eds. Teacher Skills in the 21st Century. Nobel Academic Publishing; 2022: 169-182.
4. Erdamar G, Barası M. Secondary school Turkish course curriculum in terms of 21st century skills. *J Turk Edu Sci.* 2021;19(1):312-342. doi:10.37217/tebd.868788
5. Özkaya Y, Erat V. Digital literacy studies in Turkey: literature-based qualitative research. *Nevşehir Hacı Bektaş Veli Univ J Soc Sci.* 2022;12:240-256. doi:10.30783/nevsosbilen.1124953
6. Özek Yakut B, Sincer S. Digital transformation in higher education: integration of online and hybrid practices into traditional education systems. *J Int Lang Literat Cult Histor Art Edu Res.* 2024;14:1170-1193. doi:10.51531/korkutataturkiyat.1426337
7. Compeau DR, Higgins CA. Computer self-efficacy: development of a measure and initial test. *MIS Quarterly.* 1995;19(2):189-211. doi:10.2307/249688
8. Deniz L, Algan C. Validity and reliability studies of the information and communication technologies (ICT) self-efficacy scale in education. *Marmara Univ Atatürk Fac Edu J Edu Sci.* 2013;25(25):87-107.
9. Timur B, Yılmaz Ş, Timur S. Pre-service teachers' self-efficacy beliefs towards computer use. *Mersin Univ J Fac Edu.* 2013;9(1):165-174.
10. Adıyaman M, Sert H. An investigation on teacher candidates' self-efficacy perceptions and attitudes towards computer aided education. *Akdeniz J Edu.* 2018;1(2):189-216.
11. Karaçor K. Digital Measurement Systems and Workflow. Master's thesis. İstanbul University Faculty of Dentistry; 2021.
12. Özdoğan A, Tursun B. Evaluation of dentistry graduate students' views on postgraduate education and digital dentistry. *Selcuk Dent J.* 2022; 9(1):83-87. doi:10.15311/selcukdentj.886368
13. Uyar A. Digital literacy of associate degree students. *Int J Contempor Edu Stud.* 2021;7(1):198-211.
14. Nassani LM, Bencharit S, Schumacher F, Lu WE, Resende R, Fernandes GVO. The impact of technology teaching in the dental predoctoral curriculum on students' perception of digital dentistry. *Dent J (Basel).* 2024;12(3):75. doi:10.3390/dj12030075
15. Alhamed FJ, Neiva GF, Bak SY, Karl E, Inglehart MR. Pre-doctoral dental students' computer-aided design/computer-aided manufacturing-related education, knowledge, attitudes and behavior: a national survey. *J Dent Educ.* 2023;87(4):562-571. doi:10.1002/jdd.13144
16. Liu CM, Hsu MH, Ng MY, Yu CH. Digital integration in dental education: a novel self-directed learning model using intraoral scanners for tooth preparation training. *J Dent Sci.* 2025;20(1):639-645. doi:10.1016/j.jds.2024.10.030
17. Zitzmann NU, Matthisson L, Ohla H, Joda T. Digital undergraduate education in dentistry: a systematic review. *Int J Environ Res Public Health.* 2020;17(9):3269. doi:10.3390/ijerph17093269
18. Das S, Jha K, Menon I, Kumar G, Singh A, Barman D. Application of digitalization in dental education-a systematic review. *Indian J Dent Res.* 2024;35(4):470-477. doi:10.4103/ijdr.ijdr_883_23
19. Shooriabi M, KaabOmeir F, Modarres Mousavi S, Amjadi M. Digital literacy competence among dental students: an assessment using a DigComp 2.2-based instrument. *BMC Med Educ.* 2026;26:393. doi:10.1186/s12909-026-08916-2
20. Özmen B, Ceyhan A. Virtual reality applications in dentistry. *NEU Dent J.* 2023;5(3):224-231. doi:10.51122/neudentj.2023.81
21. Tuncer M, Tanaş R. Evaluation of the computer self-efficacy perceptions of students from faculty of education. *Adıyaman Univ J Instit Soc Sci.* 2011;(6):222-232.
22. Yılmaz E, Kavak A. Evaluation of information literacy self-efficacy of university students: the case of Kırıkkale University. *Informat Manag.* 2024;7(2):348-374. doi:10.33721/by.1560442
23. Ceylan G, Eken MO, Yuruk S, Emir F. Examining the influence of self-esteem and digital literacy on professional competence factors in dental education: a cross-sectional study. *Appl Sci.* 2023;13(16):9411. doi:10.3390/app13169411

24. Erdem AR, Okul Ö. Primary school teachers' communication skills with students. *J Contempor Administrat Sci.* 2015;2(1):1-8.
25. Akyildiz S, Yurtbakan E, Tok R. The evaluation of primary school teachers' of 2023's education vision views on the politics about their professional development. *J Nat Edu Bind*. 2019;48(1):827-852.
26. Kıyıcı M. Determining the Digital Literacy Levels of Prospective Teachers. Doctoral thesis. Anadolu University; 2008.
27. Şensoy Ö. The relationship between primary school teachers' computer self-efficacy perceptions and their beliefs regarding the benefits of computer-assisted instruction. Master's thesis. Yıldız Technical University Institute of Social Sciences; 2004.
28. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Lawrence Erlbaum Associates; 1988.
29. Schwendicke F, Samek W, Krois J. Artificial Intelligence in dentistry: chances and challenges. *J Dent Res.* 2020;99(7):769-774. doi:10.1177/0022034520915714
30. Zahid TM, Agou S. E-learning readiness among dental students and faculty: a comparative study before and after the COVID-19 pandemic. *Front Med (Lausanne).* 2024;11:1306205. doi:10.3389/fmed.2024.1306205
31. Mishra P, Koehler MJ. Technological pedagogical content knowledge: a framework for teacher knowledge. *Teach Coll Record*. 2006;108(6):1017-1054. doi:10.1111/j.1467-9620.2006.00684.x
32. Koehler MJ, Mishra P. What is technological pedagogical content knowledge (TPACK)? *Contempor Issu Technol Teach Edu*. 2009;9(1):60-70.
33. Redecker C. European Framework for the Digital Competence of Educators: DigCompEdu. Publications Office of the European Union; 2017.
34. Bandura A. Self-efficacy: the exercise of control. Freeman; 1997.
35. Hatlevik OE, Guðmundsdóttir GB, Loi M. Digital diversity among upper secondary students: a multilevel analysis of the relationship between cultural capital, self-efficacy, strategic use of information and digital competence. *Comput Educ*. 2015;81:345-353. doi:10.1016/j.compedu.2014.10.019
36. Tsai M-J, Tsai C-C. Information searching strategies in web-based science learning: the role of internet self-efficacy. *Innovat Edu Teach Int*. 2003;40(1):43-50. doi:10.1080/1355800032000038822
37. Sikka N, Arya L, Bala S, et al. Knowledge, attitude, and perception of undergraduate dental students on information and communication technology and computer-assisted learning. *Int J Clin Pediatr Dent.* 2023;16(5):711-715. doi:10.5005/jp-journals-10005-2663
38. Towers A, Field J, Stokes C, Maddock S, Martin N. A scoping review of the use and application of virtual reality in pre-clinical dental education. *Br Dent J*. 2019;226(5):358-366. doi:10.1038/s41415-019-0041-0
39. Cooper J. The digital divide: the special case of gender. *J Comput Assist Learn*. 2006;22(5):320-334. doi:10.1111/j.1365-2729.2006.00185.x
40. Prensky M. Digital natives, digital immigrants part 1. *Int J Learn Futur.* 2001;9(5):1-6. doi:10.1108/10748120110424816
41. Ng W. Can we teach digital natives digital literacy? *Comput Educ*. 2012;59(3):1065-1078. doi:10.1016/j.compedu.2012.04.016
42. Helsper EJ, Eynon R. Digital natives: where is the evidence? *Br Educ Res J*. 2010;36(3):503-520. doi:10.1080/01411920902989227
43. Ferrari A. DIGCOMP: a framework for developing and understanding digital competence in Europe. Publications Office of the European Union; 2013.
44. Lave J, Wenger E. Situated learning: Legitimate peripheral participation. Cambridge University Press; 1991.
45. Kraemer-Fernandez P, Spintzyk S, Wahl E, Huettig F, Klink A. Implementation of a full digital workflow by 3D printing intraoral splints used in dental education: an exploratory observational study with respect to students' experiences. *Dent J (Basel).* 2022;11(1):5. doi:10.3390/dj11010005
46. Mwogosi A, Mambile C, Simba R, Haruna H. Integrating digital technologies in clinical dentistry training: a framework for competency-based education. *BMC Med Educ.* 2025;25(1):1645. doi:10.1186/s12909-025-07585-x
47. Gülşen İT, Yıldız Gökşen EG, Gökşen ET, Aşar EM, Meşe A. Digital dentistry among clinical-year dental students: awareness, attitudes, and curriculum expectations. *BMC Med Educ.* 2026. doi:10.1186/s12909-026-09901-5
48. Wenger E. Communities of practice: learning, meaning, and identity. Cambridge University Press; 1999.
49. van Dijk JAGM. Digital divide research, achievements and shortcomings. *Poetics.* 2006;34(4-5):221-235. doi:10.1016/j.poetic.2006.05.004
50. Bayazıt A, Ozturk Bayazit E. Digital divide affects oral and dental health students' preferences during COVID-19 outbreak: a comparison of distance and face-to-face education. *Eur Ann Dent Sci.* 2023;50(3):100-105. doi:10.52037/eads.2023.0024

Journal of Dental Sciences and Education

Review

DOI: 10.51271/JDSE-0086

Articulators in prosthodontics: evolution from mechanical to digital systems: a narrative review

Oğuzhan Kaya*, Eyyüp Altıntaş
Department of Prosthodontics, Faculty of Dentistry, Fırat University, Elazığ, Turkiye

Cite this article as: Kaya O, Altıntaş E. Articulators in prosthodontics: evolution from mechanical to digital systems: a narrative review. *J Dent Sci Educ.* 2026;4(3):119-123. doi:10.51271/JDSE-0086

Received: 12.10.2025 ◆ **Accepted**: 27.07.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Articulators are mechanical or digital devices that simulate mandibular movements and reproduce maxillomandibular relationships in prosthodontic procedures. Throughout prosthodontic history, articulators have evolved from simple hinge devices to fully adjustable mechanical and virtual systems. This narrative review summarizes conventional and digital articulator systems, with emphasis on their clinical indications, limitations, and relevance to contemporary prosthodontic workflows. Mechanical articulators remain useful for diagnostic mounting, treatment planning, and laboratory communication; however, they cannot fully reproduce patient-specific neuromuscular control, tissue resilience, tooth mobility, or dynamic occlusion. Digital facebows, mandibular motion analyzers, computer-aided design/computer-aided manufacturing (CAD/CAM) systems, and virtual articulators may improve data integration and visualization, but their clinical value depends on the quality of input records, software compatibility, cost, and clinician experience. A critical understanding of each system is therefore necessary for selecting an articulator according to treatment complexity, occlusal risk, vertical dimension changes, and the need for individualized mandibular movement simulation.

Keywords: Dental occlusion, mandibular movement, temporomandibular joint, jaw relation record, computer-aided design, computer-aided manufacturing

INTRODUCTION

Articulators are mechanical devices designed to mimic all or part of jaw movements, to which upper and lower models can be attached, and which can represent the jaws and temporomandibular joint (TMJ).[1] They are used for diagnosis, treatment planning, prosthesis fabrication, and laboratory communication. For accurate prosthodontic treatment, the movements and positional relationships of the mandible must be transferred and interpreted as reliably as possible.[2]

Although articulators may reproduce selected anatomical and functional determinants of mandibular movement, they should not be considered exact replicas of the stomatognathic system. Muscles, ligaments, the TMJ, bone, connective tissues, neuromuscular adaptation, tissue resilience, and tooth mobility cannot be fully reproduced in mechanical devices.[3,4] Therefore, the value of an articulator depends not only on its mechanical complexity but also on the accuracy of clinical records and the clinician's ability to interpret its limitations.

Historically, before the first mechanical articulator attributed to Gariot in 1805, plaster occludators were used to preserve horizontal and vertical relationships between casts.[5] Since then, articulators have evolved from simple hinge devices to semi-adjustable, fully adjustable, and virtual systems. The current clinical question is not whether one articulator is universally superior, but which system provides sufficient accuracy and practicality for a specific treatment scenario.

CLINICAL ROLE AND INHERENT LIMITATIONS OF CONVENTIONAL ARTICULATORS

The clinical value of conventional articulators lies in transferring maxillomandibular relationships to a stable laboratory setting, reducing chairside adjustments, and improving communication between the clinician and the dental technician. Nevertheless, their limitations must be interpreted critically. Increasing mechanical complexity does not automatically improve clinical outcome if facebow transfer, interocclusal records, cast mounting, or occlusal analysis are inaccurate.[6-8] These clinical benefits and limitations are summarized in **Table 1**.

ARTICULATOR RECORDS AND CLASSIFICATION

Traditional classifications of articulators are based on the type of mandibular movement reproduced, the records accepted by the device, and the degree of adjustability. Instead of presenting long historical lists of device names, the clinically relevant distinction is whether the articulator can accept facebow transfer, interocclusal records, graphic records, and patient-specific movement information.[9-18] A clinically oriented classification is presented in **Table 2**.

*Corresponding Author: Oğuzhan Kaya, o.kaya@firat.edu.tr

Table 1. Clinical value and limitations of conventional articulators

Clinical aspect	Potential benefit	Main limitation	Clinical implication
Diagnostic mounting	Permits extraoral evaluation of occlusion and jaw relations	Static casts cannot reproduce neuromuscular adaptation or tissue resilience	Useful for planning, but intraoral verification remains necessary
Laboratory procedures	Allows occlusal adjustments and wax-up/restoration design without the patient present	Errors may arise from record deformation, mounting inaccuracies, and articulator instability	Accurate records and careful mounting are as important as articulator type
Mandibular movement simulation	Higher-level articulators can approximate condylar guidance, Bennett movement, and intercondylar distance	Mechanical pathways remain simplified compared with patient-specific functional movement	Semi- or fully adjustable devices are more justified in extensive rehabilitations than in simple single-unit treatments
Clinical efficiency	May reduce chairside occlusal adjustments when the selected system matches the clinical problem	Complex systems require additional time, cost, and operator skill	Selection should balance accuracy, practicality, treatment extent, and clinician experience

Table 2. Clinically oriented classification of conventional articulators

Articulator category	Records commonly required	Strengths	Limitations/appropriate use
Simple hinge articulators	Centric relation or maximum intercuspation record only	Inexpensive, rapid mounting, suitable for limited procedures when eccentric contacts are not being developed	Cannot orient casts to the TMJ or simulate eccentric movements; not appropriate for complex occlusal rehabilitation
Mean-value articulators	Average condylar/incisal settings; usually no individualized dynamic records	More functional than hinge articulators and practical for simple prosthetic procedures	Uses population-based values; limited individualization and limited relevance when anterior guidance is absent or vertical dimension is changed
Semi-adjustable articulators	Facebow transfer and protrusive/lateral interocclusal records	Allows adjustment of condylar inclination, Bennett angle, and selected guidance parameters; useful for routine fixed/removable prosthodontics	Movements are still mechanically simplified; record quality strongly influences accuracy
Fully adjustable articulators	Kinematic facebow, pantographic/graphic records, and interocclusal records	Allows the highest level of individualization among mechanical systems	Technique-sensitive, time-consuming, and rarely necessary for simple cases; best reserved for complex occlusal rehabilitation
Arcon/non-arcon design	Depends on the chosen category and mounting protocol	Arcon design resembles the anatomical relationship of condyle and fossa; non-arcon designs may provide simpler centric locking	Evidence is not uniform regarding clinical superiority; the operator must understand how design affects vertical dimension changes and lateral movement interpretation

TMJ: Temporomandibular joint

ARCON AND NON-ARCON DESIGNS

Arcon and non-arcon articulators differ in the location of the mechanical condylar and fossae elements. In arcon articulators, the mechanical fossae are located on the upper member, approximating the anatomical relationship of the glenoid fossae to the cranium. In non-arcon systems, the arrangement is reversed.[11,18] The literature does not support a simple assumption that one design is always superior; rather, the clinical effect depends on record accuracy, vertical dimension changes, and the intended restorative procedure. Goyal et al.[19] reported different sagittal condylar guidance values between arcon and non-arcon articulators, whereas previous reports have also noted comparable condylar guidance or improved mandibular movement representation with arcon devices.[18-20] The main clinical differences between the two designs are summarized in **Table 3**.

CLINICAL SELECTION OF ARTICULATORS

Articulator selection should be guided by four variables: the patient's occlusal characteristics, the extent of restorative treatment, the limitations of the device, and the clinician's skill.[16,21,22] Simple or mean-value articulators may be acceptable when restoring a single unit or when stable anterior guidance is present. Semi-adjustable articulators are generally more appropriate for routine fixed prosthodontics, balanced complete denture arrangements, and removable partial dentures. Fully adjustable articulators are most justified when extensive fixed prostheses, full-mouth rehabilitation, major occlusal adjustment, or vertical dimension changes are planned. The most sophisticated device may still lead to inaccurate outcomes if clinical records are poor; conversely, a semi-adjustable articulator used by an experienced clinician may outperform a fully adjustable device used without proper records.

FACEBOW RECORDS: EVIDENCE AND CLINICAL INDICATIONS

A facebow records the spatial relationship of the maxillary arch to selected cranial reference points and transfers this relationship to the articulator.[1] Although prosthodontic

Table 3. Comparative clinical interpretation of arcon and non-arcon articulators

Design feature	Arcon articulator	Non-arcon articulator	Clinical interpretation
Anatomic analogy	Condylar elements are on the lower member and fossae on the upper member	Condylar elements are generally on the upper member	Arcon design is more intuitive for demonstrating mandibular movements
Vertical dimension changes	Condylar guidance settings are less affected by opening the articulator	Changes in vertical dimension may require greater caution or readjustment	Arcon systems may be advantageous when vertical dimension is altered
Laboratory handling	Upper and lower members can often be separated easily	More compact assembly and simpler centric locking	Ease of laboratory handling must be balanced against maintenance of centric relation
Evidence-based selection	Some studies report more accurate movement representation	Some reports suggest similar condylar guidance values	Selection should not rely on design alone; records, calibration, and clinician skill are decisive

teaching has traditionally emphasized facebow use, clinical evidence is mixed.[23-25] In extensive fixed restorations, full-mouth rehabilitation, or cases involving alteration of vertical dimension, facebow records may be more defensible because errors in the arc of closure may be clinically relevant.[26,27] For complete dentures and occlusal splints, systematic review evidence has not shown consistent superiority of facebow transfer.[23,28-30] The main facebow-related records and clinical considerations are summarized in **Table 4**.

in a digital environment.[40-42] It should be regarded as an adjunctive diagnostic and planning tool rather than a complete replacement for biological assessment. Virtual articulators may incorporate average-value settings, semi-adjustable mechanical parameters, or patient-specific movement data from jaw tracking systems.[40,43-46] Their clinical utility is greatest when data acquisition and transfer are standardized. The principal data inputs and limitations of virtual approaches are summarized in **Table 5**.

Table 4. Facebow records and evidence-based clinical use

Record/system	Clinical purpose	Advantages	Limitations/evidence considerations
Arbitrary facebow	Transfers the maxillary cast to an articulator using average hinge-axis references	Simple, widely taught, and useful for many semi-adjustable articulator protocols	Skin and external reference points may be unreliable; benefit for complete dentures and splints is not consistently supported
Kinematic facebow	Locates the transverse horizontal hinge axis more precisely	Provides more individualized axis transfer for fully adjustable systems	Time-consuming and technique-sensitive; may be uncomfortable or unstable in edentulous patients
Digital facebow	Aligns facial, intraoral, and/or jaw-motion data in a digital workflow	Improves data storage, transfer, visualization, and clinician-technician communication	Requires scanners, software, calibration, and interoperable data formats
No facebow/simplified mounting	Uses average mounting or simplified jaw relation records	May be practical for simple cases and complete dentures where evidence does not show clear superiority of facebow use	Less suitable when vertical dimension is altered or full-arch fixed prostheses require individualized transfer

PANTOGRAPHIC AND DIGITAL RECORDS

Pantographs graphically record mandibular movement pathways and provide information for articulator adjustment.[31] In contemporary workflows, digital facebows and jaw-motion tracking systems extend this concept by integrating facial scans, intraoral scans, cone-beam computed tomography (CBCT) data, and mandibular movement records.[32-39] Digitalization may reduce physical record storage, improve repeatability, and facilitate communication; however, accuracy is dependent on scanner calibration, registration protocols, software alignment, and operator training.

Digital facebow techniques share a common principle: digital facial data are aligned with intraoral or cast scan data to determine the spatial position of the maxillary arch. Systems and protocols vary, including the use of optical face scanners, jaw tracking devices, intraoral scanners, photogrammetry, and virtual mounting software.[34,37-39] Therefore, current evidence should be interpreted according to the specific hardware/software combination rather than generalized to all digital facebow systems.

VIRTUAL ARTICULATORS AND DIGITAL OCCLUSAL ANALYSIS

A virtual articulator is software that simulates maxillomandibular relationships and mandibular movements

Earlier digital articulator designs and integrated craniofacial movement models also demonstrate the broad range of virtual simulation approaches.[47-53]

Several studies suggest that full-arch digital impressions and virtual interocclusal registration may provide clinically useful data when appropriate software and scanning protocols are used.[54,55] However, virtual contacts should be interpreted cautiously because scanner accuracy, matching algorithms, mandibular recording errors, and data export formats can affect the final occlusal design. The lack of a universal digital format for mandibular movement transfer remains a relevant limitation.[49]

CRITICAL COMPARISON OF CONVENTIONAL AND DIGITAL SYSTEMS

The transition from mechanical to virtual articulators should be understood as an expansion of available clinical tools rather than a linear replacement. Mechanical articulators are accessible, familiar, and adequate for many treatments, whereas virtual systems can improve visualization, dynamic occlusal assessment, and digital design integration. Yet digital systems are not free of error; they shift the source of error from stone casts and mechanical joints to scan accuracy, registration algorithms, software compatibility, and clinician training. In addition, adoption of fully digital workflows can be influenced by the clinician's learning curve.[56] The major comparative advantages and limitations of conventional and digital systems are summarized in **Table 6**.

Table 5. Digital and virtual articulator approaches

Virtual approach/system	Data input	Potential benefit	Main limitation
Mathematically simulated articulator	Average or user-defined condylar and incisal parameters	Easy integration with CAD/CAM design and educational visualization	Relies on average values; cannot fully represent individual mandibular dynamics
Patient-specific virtual articulator	Jaw movement analyzer, pantographic/dynamic records, intraoral scans, and facial/CBCT data when available	Allows static and dynamic occlusal contact analysis during mandibular movement	Requires additional hardware, calibration, and technical expertise
Optical systems such as Modjaw or Freecorder BlueFox	Optical tracking of mandibular movement and digital model integration	Can visualize functional movement and support dynamic occlusal mapping	Cost and interoperability vary across systems
CBCT/4D-integrated systems	Radiographic data combined with facial or mandibular tracking data	May support three-dimensional visualization of craniofacial structures and jaw motion	Radiation exposure, hardware requirements, and clinical indications must be justified

CBCT: Cone-beam computed tomography, CAD/CAM: Computer-aided design/computer-aided manufacturing

Table 6. Critical comparison of conventional and digital articulator systems

System	Main advantages	Main limitations	Most relevant indications
Mechanical articulators	Low to moderate cost, familiar protocols, physical handling of casts, compatibility with many laboratory workflows	Limited simulation of dynamic occlusion, potential mounting and material errors, inability to reproduce tissue resilience and tooth mobility	Diagnostic mounting, routine fixed/removable prosthodontics, complete denture tooth arrangement
Semi-adjustable articulators	Practical balance between individualization and efficiency; accepts facebow and eccentric records	Still mechanically simplified; depends heavily on record accuracy	Routine crown/bridge work, removable prostheses, balanced occlusion arrangements
Fully adjustable mechanical articulators	Highest mechanical individualization and compatibility with pantographic records	Complex, time-consuming, costly, and technique-sensitive	Full-mouth rehabilitation, major vertical dimension alteration, complex occlusal analysis
Virtual articulators	Digital storage, visualization, CAD/CAM integration, potential dynamic occlusal analysis, improved communication	Software/hardware cost, interoperability issues, data-entry requirements, dependence on scanner and motion-recording accuracy	Digital prosthodontics, full-arch digital workflows, complex treatment planning, patient education

CAD/CAM: Computer-aided design/computer-aided manufacturing

CONCLUSION

Articulators remain important instruments in prosthodontics because they allow the clinician and technician to evaluate maxillomandibular relationships outside the oral cavity. However, the clinical value of an articulator is determined by the match between device capability, treatment complexity, record accuracy, and clinician experience. Mechanical systems remain suitable for many conventional procedures, while digital facebows, jaw motion analyzers, and virtual articulators provide additional opportunities for three-dimensional visualization and individualized digital restoration design. Current evidence supports a selective rather than universal approach. Simple or mean-value articulators may be adequate for limited procedures, semi-adjustable devices provide a practical standard for routine prosthodontics, and fully adjustable or virtual systems are most justified in complex rehabilitations, vertical dimension changes, and workflows requiring dynamic occlusal analysis. Future improvements in sensor accuracy, software interoperability, and standardized digital record transfer may further increase the role of virtual articulators in routine prosthodontic practice.

ETHICAL DECLARATIONS

Peer Review Process

This manuscript was subject to external peer review.

Conflict of Interest

The authors declare no conflicts of interest related to this study.

Financial Disclosure

The authors received no financial support for the conduct or publication of this research.

Author Contributions

Concept: OK, EA; Design: OK, EA; Control: OK, EA; Resources: OK, EA; Materials: OK, EA; Data Collection and/or Processing: OK, EA; Analysis and/or Interpretation: OK, EA; Literature Review: OK, EA; Writing the Article: OK, EA; Critical Review: OK, EA.

REFERENCES

1. The glossary of prosthodontic terms: ninth edition. *J Prosthet Dent.* 2017;117(5S):e1-e105. doi:10.1016/j.prosdent.2016.12.001
2. Stuart CE. Accuracy in measuring functional dimensions and relations in oral prosthesis. *J Prosthet Dent.* 1959;9(2):220-236. doi:10.1016/0022-3913(59)90007-1
3. de Carvalho ODT. A new fully adjustable articulator system and procedure. *J Prosthet Dent.* 1998;80(3):376-386. doi:10.1016/S0022-3913(98)70141-6
4. Çalıkkocaoğlu S. Dişsiz Hastaların Protetik Tedavisi: Klasik Tam Protezler. 5th ed. Quintessence Yayıncılık; 2010.
5. Çalıkkocaoğlu S. *Tam Protezler.* Vol 1. 4th ed. Protez Akademisi ve Gnatoloji Derneği; 2004:329-395.
6. Rosenstiel SF, Land MF, Fujimoto J. Contemporary Fixed Prosthodontics. 5th ed. Elsevier/Mosby; 2015:110-145.
7. Nallaswamy D. Textbook of Prosthodontics. Jaypee Brothers Medical Publishers; 2003:153-166.
8. Heartwell CM Jr, Rahn AO. Syllabus of Complete Dentures. 2nd ed. Lea & Febiger; 1974.
9. Kulunk S. Artikülatörler ve İnteroklüzal Kayıt. Lecture notes. Ondokuz Mayıs University Faculty of Dentistry. Accessed June 16, 2026.
10. Boucher CO, Hickey JC, Zarb GA, Swenson MG. Prosthodontic Treatment for Edentulous Patients. 7th ed. CV Mosby; 1975.
11. Çalıkkocaoğlu S. Tam Protezler. Vol 1. 4th ed. Protez Akademisi ve Gnatoloji Derneği; 2004:329-395.
12. Ergin U. Toplumumuzun bir kesiminden elde edilen ölçümlere uygun olarak deneysel bir artikülatör modelinin oluşturulması. Doctoral Dissertation. 1999:12-13,16-23,24-29,30-69,71-97.
13. Çalıkkocaoğlu S. Tam Protezler. Vol 1. 3rd ed. Protez Akademisi ve Gnatoloji Derneği; 1998:329-362.
14. Starcke EN. The history of articulators: early attempts to reproduce mandibular movement, part II. *J Prosthodont.* 2000;9(2):110-112. doi:10.1111/j.1532-849X.2000.00110.x
15. Sarandha DL. Textbook of Complete Denture Prosthodontics. Jaypee Brothers Medical Publishers; 2008.
16. Okeson JP. Management of Temporomandibular Disorders and Occlusion. 5th ed. Mosby; 2003.
17. Shillingburg HT, Hobo S, Whitsett LD. Fundamentals of Fixed Prosthodontics. 2nd ed. Quintessence Publishing; 1981:115-117.
18. Tylman SD. Theory and Practice of Crown and Fixed Partial Prosthodontics (Bridge). 6th ed. Mosby; 1970.
19. Goyal MK, Goyal S. A comparative study to evaluate the discrepancy in condylar guidance values between two commercially available arcon and non-arcon articulators: a clinical study. *Indian J Dent Res.* 2011;22(6):880-886. doi:10.4103/0970-9290.94695
20. Beck HO. A clinical evaluation of the arcon concept of articulation. *J Prosthet Dent.* 1959;9(3):409-421. doi:10.1016/0022-3913(59)90072-1
21. Hobo S, Shillingburg HT Jr, Whitsett LD. Articulator selection for restorative dentistry. *J Prosthet Dent.* 1976;36(1):35-43. doi:10.1016/0022-3913(76)90231-6
22. Zarb GA, Bolender CL, Eckert SE, Jacob RF, Fenton AH, Mericske-Stern R. Prosthodontic Treatment for Edentulous Patients: Complete Dentures and Implant-Supported Prostheses. 12th ed. Mosby; 2004.
23. Khan FR, Ali R, Sheikh A. Utility of facebow in the fabrication of complete dentures, occlusal splints and full arch fixed dental prostheses: a systematic review. *Indian J Dent Res.* 2018;29(3):364-369. doi:10.4103/ijdr.IJDR_377_17
24. Lynch CD, Singhrao H, Addy LD, Gilmour ASM. The teaching of fixed partial dentures in undergraduate dental schools in Ireland and the United Kingdom. *J Oral Rehabil.* 2010;37(12):908-915. doi:10.1111/j.1365-2842.2010.02106.x

25. Collet HA. The movements of temporomandibular joints and their relation to the problems of occlusion. *J Prosthet Dent.* 1955;5:486-496.
26. Johnston JF, Phillips RW, Dykema RW. Kuron ve Köprü Protezlerinde Modern Uygulamalar. Dr İbrahim Çağlayan Mezuniyet Sonrası Eğitimi ve Bilimsel Teknik Araştırma Vakfı; 1982.
27. Squier RS. Jaw relation records for fixed prosthodontics. *Dent Clin North Am.* 2004;48(2):471-486. doi:10.1016/j.cden.2004.01.001
28. Wang MQ, Xue F, Chen J, Fu K, Cao Y, Raustia A. Evaluation of the use of and attitudes towards a face-bow in complete denture fabrication: a pilot questionnaire investigation in Chinese prosthodontists. *J Oral Rehabil.* 2008;35(9):677-681. doi:10.1111/j.1365-2842.2007.01835.x
29. Fox SS. The significance of errors in hinge axis location. *J Am Dent Assoc.* 1967;74(6):1268-1272. doi:10.14219/jada.archive.1967.0448
30. Ellinger CW, Wesley RC, Abadi BJ, Armentrout TM. Patient response to variations in denture technique. Part VII: twenty-year patient status. *J Prosthet Dent.* 1989;62(1):45-48. doi:10.1016/0022-3913(89)90045-0
31. Brewka RE. Pantographic evaluation of cephalometric hinge axis. *Am J Orthod.* 1981;79(1):1-19. doi:10.1016/0002-9416(81)90097-X
32. Tokgöz Çetindağ M, Meşe A. Diş hekimliğinde kullanılan CAD/CAM (bilgisayar destekli tasarım/bilgisayar destekli üretim) sistemleri ve materyaller. *Atatürk Üniv Diş Hek Fak Derg.* 2016;26(3):524-533. doi:10.17567/ataunidfd.290443
33. Janeva NM, Kovacevska G, Elencevski S, Panchevska S, Mijoska A, Lazarevska B. Advantages of CAD/CAM versus conventional complete dentures-a review. *Open Access Maced J Med Sci.* 2018;6(8):1498-1502. doi:10.3889/oamjms.2018.308
34. Çölgeçen Ö, Köse M. Protetik diş hekimliğinde dijital yüz arkları. *Int Arch Dent Sci.* 2021;42(1):65-69. doi:10.5505/eudfd.2021.78790
35. Shah K, Koka S. Evidence-based practice and barriers to compliance: face bow transfer. *J Prosthodont Res.* 2016;60(1):20-22. doi:10.1016/j.jpor.2015.09.004
36. Shodadai SP, Türp JC, Gerds T, Strub JR. Is there a benefit of using an arbitrary facebow for the fabrication of a stabilization appliance? *Int J Prosthodont.* 2001;14(6):517-522.
37. Solaberrieta E, Minguez R, Barrenetxea L, Etxaniz O. Direct transfer of the position of digitized casts to a virtual articulator. *J Prosthet Dent.* 2013;109(6):411-414. doi:10.1016/S0022-3913(13)60330-3
38. Lam WYH, Hsung RTC, Choi WWS, Luk HWK, Pow EHN. A 2-part facebow for CAD-CAM dentistry. *J Prosthet Dent.* 2016;116(6):843-847. doi:10.1016/j.prosdent.2016.05.013
39. Kalman L, Chrapka J, Joseph Y. Digitizing the facebow: a clinician/technician communication tool. *Int J Prosthodont.* 2016;29(1):35-37. doi:10.11607/ijp.4748
40. Lepidi L, Galli M, Mastrangelo F, et al. Virtual articulators and virtual mounting procedures: where do we stand? *J Prosthodont.* 2021;30(1):24-35. doi:10.1111/jopr.13240
41. Maestre-Ferrín L, Romero-Millán J, Peñarrocha-Oltra D, Peñarrocha-Diago M. Virtual articulator for the analysis of dental occlusion: an update. *Med Oral Patol Oral Cir Bucal.* 2012;17(1):e160-e163. doi:10.4317/medoral.17147
42. Özdemir G, Albayrak B, Yüzbaşıoğlu E, Ölçer Us Y. Virtual articulators, virtual occlusal records and virtual patients in dentistry. *J Exp Clin Med.* 2021;38(3S):129-135. doi:10.52142/OMUJECM.38.SI.DENT.9
43. Szentpétery A. Dynamic correction of occlusal surfaces by means of the CAD-CAM methods. II. *Fogorv Sz.* 1999;92(9):263-272.
44. Bisler A, Bockholt U, Voss G. The virtual articulator-applying VR technologies to dentistry. In: Proceedings of the Sixth International Conference on Information Visualisation. IEEE; 2002:600-602. doi:10.1109/IV.2002.1028835
45. Wakabayashi K, Sohmura T, Takahashi J, et al. Development of the computerized dental cast form analyzing system: three-dimensional diagnosis of dental arch form and investigation of measuring condition. *Dent Mater J.* 1997;16(2):180-190,225. doi:10.4012/dmj.16.180
46. Kordass B, Gärtner C, Söhnel A, et al. The virtual articulator in dentistry: concept and development. *Dent Clin North Am.* 2002;46(3):493-506. doi:10.1016/S0011-8532(02)00006-X
47. Solaberrieta E, Etxaniz O, Minguez R, Muniozguren J, Arias A. Design of a virtual articulator for the simulation and analysis of mandibular movements in dental CAD/CAM. In: Proceedings of the 19th CIRP Design Conference-Competitive Design. Cranfield University Press; 2009:323.
48. Koralakunte PR, Aljanakh M. The role of virtual articulator in prosthetic and restorative dentistry. *J Clin Diagn Res.* 2014;8(7):ZE25-ZE28. doi:10.7860/JCDR/2014/8929.4648
49. Park S. Digitalization of Virtual Articulator: Methods, Discrepancy to Real Articulators, Comparing of Each Methods. Master's thesis. Lithuanian University of Health Sciences; 2017.
50. Enciso R, Memon A, Mah J. Three-dimensional visualization of the craniofacial patient: volume segmentation, data integration and animation. *Orthod Craniofac Res.* 2003;6(suppl 1):66-71. doi:10.1034/j.1600-0544.2003.237.x
51. Solaberrieta E, Barrenetxea L, Minguez R, Iturrate M, De Prado I. Registration of mandibular movement for dental diagnosis, planning and treatment. *Int J Interact Des Manuf.* 2018;12(3):1027-1038. doi:10.1007/s12008-017-0438-4
52. Dai F, Wang L, Chen G, Chen S, Xu T. Three-dimensional modeling of an individualized functional masticatory system and bite force analysis with an orthodontic bite plate. *Int J Comput Assist Radiol Surg.* 2016;11(2):217-229. doi:10.1007/s11548-015-1248-4
53. Lippold C, Hoppe G, Moiseenko T, Ehmer U, Danesh G. Analysis of condylar differences in functional unilateral posterior crossbite during early treatment: a randomized clinical study. *J Orofac Orthop.* 2008;69(4):283-296. doi:10.1007/s00056-008-0803-9
54. Keul C, Güth JF. Accuracy of full-arch digital impressions: an in vitro and in vivo comparison. *Clin Oral Investig.* 2020;24(2):735-745. doi:10.1007/s00784-019-02965-2
55. Edher F, Hannam AG, Tobias DL, Wyatt CCL. The accuracy of virtual interocclusal registration during intraoral scanning. *J Prosthet Dent.* 2018;120(6):904-912. doi:10.1016/j.prosdent.2018.01.024
56. Agnini A, Agnini AM, Coachman C. Digital Dental Revolution: The Learning Curve. Quintessence Publishing; 2015.

Journal of Dental Sciences and Education

Case Report

DOI: 10.51271/JDSE-0087

Surgery for widespread bone and soft tissue necrosis associated with extrusion of a calcium hydroxide paste containing barium sulfate: a case report with submandibular spread

Halil İbrahim Durmuş*, Mehmet Sait Şimşek, Tevfik Kızılseki, Kağan Fatih Kuşan

Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Adıyaman University, Adıyaman, Turkiye

Cite this article as: Durmuş Hİ, Şimşek MS, Kızılseki T, Kuşan KF. Surgery for widespread bone and soft tissue necrosis associated with extrusion of a calcium hydroxide paste containing barium sulfate: a case report with submandibular spread. *J Dent Sci Educ.* 2026;4(3):124-127. doi:10.51271/JDSE-0087

Received: 08.04.2025 ◆ **Accepted**: 31.07.2026 ◆ **Published**: 20.08.2026

ABSTRACT

Calcium hydroxide, a commonly used intracanal medication in endodontic treatment, can lead to serious bone and soft tissue damage if it is extruded uncontrollably. This case report describes the surgical management of extensive tissue necrosis resulting from the extrusion of a calcium hydroxide paste containing barium sulfate. A fourteen-year-old female patient presented with severe pain and diffuse, firm swelling in the posterior right mandibular region after root canal treatment. Cone beam computed tomography revealed that the extruded material extended beyond the lingual cortex into the submandibular space, causing bone destruction; however, the inferior alveolar nerve was spared. Surgical treatment included removal of the extruded material, debridement of necrotic tissue, and extraction of the affected tooth. Histopathological examination revealed necrosis and a foreign body reaction. No complications occurred during the postoperative period. This case demonstrates that serious tissue destruction can occur even without nerve damage and highlights the importance of early surgical interventions.

Keywords: Calcium hydroxide extrusion, necrotic changes in bone tissue, complications related to endodontic treatment, surgical debridement, inferior alveolar nerve

INTRODUCTION

Calcium hydroxide is an intracanal medication widely used in endodontic treatments because of its potent antimicrobial properties and high alkaline pH.[1,2] It is frequently preferred in clinical practice because it reduces the microbial load in the root canal system and contributes to periapical tissue healing. However, if the material extends beyond the apical limits, it can cause serious and, in some cases, irreversible damage to the surrounding hard and soft tissues.[3]

The high pH level of calcium hydroxide (approximately 12.5) causes tissue necrosis through protein denaturation, disruption of cell membrane integrity, and adverse effects on microvascular circulation. This pathophysiological process may manifest as pain, inflammation, ulceration, and delayed healing.[4] The literature reports that such complications are most frequently associated with inferior alveolar nerve injury, with neurosensory disorders such as paresthesia, dysesthesia, and anesthesia being among the primary symptoms.[5]

However, current case reports indicate that significant bone and soft tissue destruction can occur, despite the preservation of nerve structures.[6] The potential for extruded material to advance through anatomical spaces is considered a determining factor in the extent of damage. The thinness of the lingual cortex, particularly in the posterior mandibular region, can facilitate the passage of materials into the submandibular space. This study aims to present the surgical treatment of bone and soft tissue necrosis extending to the submandibular space due to extrusion of a calcium hydroxide paste containing barium sulfate and to highlight the importance of early diagnosis and an appropriate surgical approach in such complications.

CASE

A 14-year-old girl in good general health presented to our clinic with symptoms that developed following the application of intracanal calcium hydroxide during an interim session of root canal treatment initiated on the right mandibular second molar (#47). The patient reported intense pain in the right mandibular region that began shortly after the procedure and increased in severity over time, accompanied by firm and widespread swelling. The symptoms increased significantly within the first 24-48 h, and sensitivity during chewing and palpation became more pronounced (**Figure 1**). No findings suggestive of fever, systemic illness, or systemic infection were observed in the patient. Panoramic radiography revealed a radiopaque appearance around the affected tooth, consistent with extruded material (**Figure 2**). Cone-beam computed tomography revealed that the foreign

***Corresponding Author:** Halil İbrahim Durmuş, halilibrahimdurmus@adiyaman.edu.tr

body had penetrated the lingual cortical bone, causing bone destruction in the mandibular basal region and extending into the submandibular space (**Figure 3**). However, the integrity of the inferior alveolar nerve canal was maintained. Clinical neurosensory examination was performed using light-touch, pin-prick, and two-point discrimination tests in the distribution of the inferior alveolar and mental nerves. No sensory deficits were detected at presentation. Based on the clinical and radiographic findings, a preliminary diagnosis of chemical tissue injury and osteonecrotic changes secondary to extrusion of a calcium hydroxide paste containing barium sulfate was made. Owing to the progressive nature of the clinical picture, surgical treatment was planned. Under local anesthesia, a lingual sulcular incision was made, and a full-thickness mucoperiosteal flap was elevated (**Figure 4**). Local anesthesia was achieved using 4% articaine with 1:100,000 epinephrine. Following flap elevation, extensive necrotic soft tissue and devitalized bone were identified. The affected soft tissues appeared grayish, nonviable, and lacked normal vascularity, consistent with extensive soft tissue necrosis. The clinical appearance was consistent with extensive soft tissue necrosis, which was subsequently confirmed by histopathological examination. The extruded calcium hydroxide paste was carefully removed using surgical curettes and copious saline irrigation. Debridement was continued until viable bleeding bone and healthy soft tissue margins were observed.

Figure 1. Firm swelling of the right mandible with facial asymmetry secondary to paste extrusion

Figure 2. Radiopaque lesion in the right mandible consistent with extruded paste

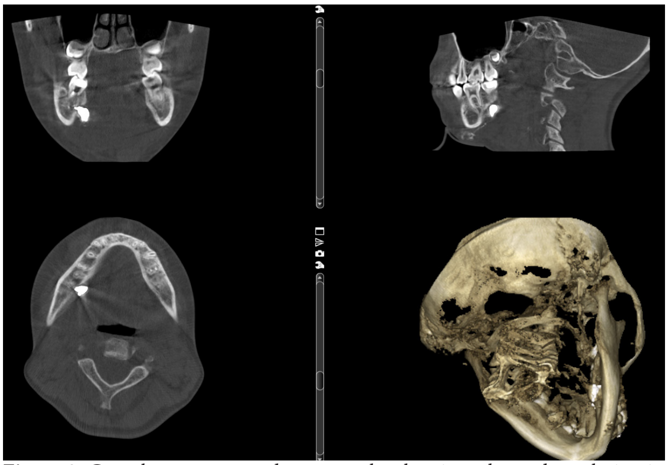

Figure 3. Cone-beam computed tomography showing a hyperdense lesion in the right mandible with surrounding bone necrosis

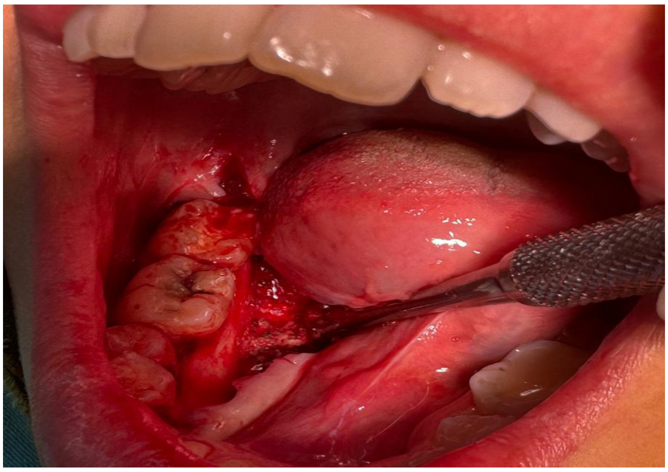

Figure 4. Intraoperative view of extruded paste following the sulcular incision

During surgical exploration, necrotic soft tissue and destructive bone were observed. The extruded material was carefully removed, and the necrotic tissue was debrided (**Figure 5**). Tooth 47, which had a poor prognosis, was extracted. Extraction was considered necessary because of extensive surrounding bone destruction, contamination by the extruded material, compromised periodontal support, and the low likelihood of predictable long-term functional preservation. The surgical site was copiously irrigated with sterile saline and rifampicin solution to remove residual foreign material and necrotic debris before primary closure was achieved with sutures. Histopathological examination revealed necrotic bone tissue, inflammatory cell infiltration, and foreign body reactions. These findings support the clinical diagnosis.

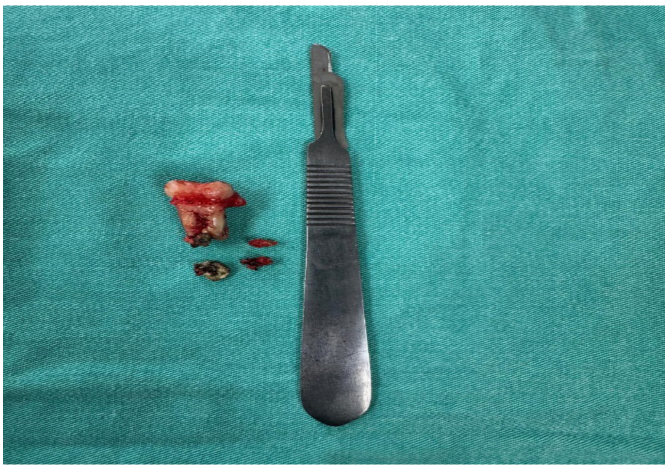

Figure 5. Extracted tooth No. 47 with the paste material removed

In the postoperative period, the patient received amoxicillin-clavulanic acid, nonsteroidal anti-inflammatory analgesics, and chlorhexidine mouthwash. Detailed oral hygiene instructions were provided, and the patient was closely monitored for signs of infection, wound dehiscence, or neurosensory complications.

The patient was followed up regularly, and no complications were observed at the 1-week and 1-month follow-up visits (**Figure 6, 7**).

Figure 6. Panoramic radiograph taken 1 week postoperatively showing removal of the paste material

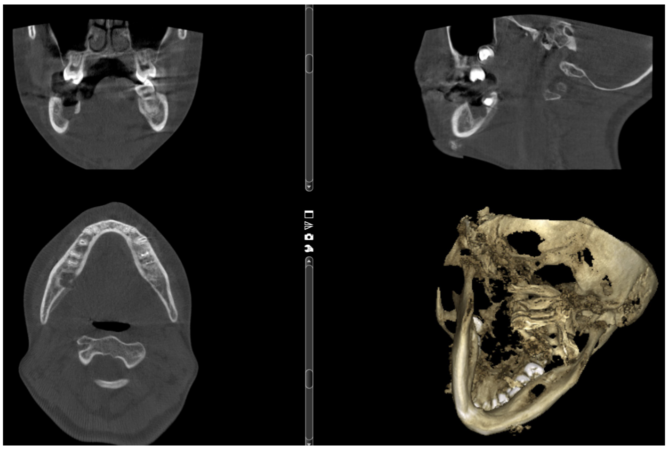

Figure 7. Cone-beam computed tomography at 1 month postoperatively demonstrated bone healing

DISCUSSION

Although rare, calcium hydroxide extrusion can lead to serious iatrogenic complications. The chemical burn effect, resulting from a highly alkaline pH, causes coagulation necrosis through protein denaturation and microvascular damage.[4] Calcium hydroxide-induced tissue injury is mediated by multiple mechanisms, including cellular membrane destruction, vascular thrombosis, protein denaturation, and disruption of local blood supply. These effects may lead to progressive soft tissue necrosis and osteonecrosis, particularly when large volumes of material are extruded into confined anatomical spaces. In the literature, the vast majority of these complications have been associated with inferior alveolar nerve injury.[5] However, it is noteworthy that the nerve structures were preserved in the present case.

Previously reported cases have generally described sensory disturbances following extrusion into the inferior alveolar canal, frequently requiring prolonged follow-up and, in some instances, microsurgical intervention. Unlike those reports, the present case was characterized by extensive osseous and soft tissue necrosis without neurosensory impairment, suggesting that tissue destruction may occur independently of direct nerve injury.

Nevertheless, the significant bone destruction and submandibular extension observed indicate that tissue damage is not limited to neurological findings.[6] This highlights the need for clinicians to consider that serious complications can develop even without neurological findings. The ability of the extruded material to spread through anatomical spaces facilitates its extension into the submandibular space, especially in the presence of lingual cortical perforation. This suggests that mechanical spread, along with chemical toxicity, plays a role in pathogenesis. Preservation of the inferior alveolar nerve in this patient may be explained by the anatomical pathway of material dissemination. CBCT imaging indicated preferential spread through the lingual cortex toward the submandibular space, thereby minimizing direct contact with the neurovascular bundle. Although the exact mechanism of extrusion could not be determined retrospectively, excessive pressure during intracanal medicament placement and the presence of an open apical pathway may have contributed to displacement of the calcium hydroxide paste beyond the root canal system. Similar mechanisms have been proposed in previously reported cases of calcium hydroxide extrusion.[7]

Cone-beam computed tomography plays a significant role in the diagnosis of this condition. Three-dimensional imaging allows for detailed assessment of the location of the extruded material, the direction of its spread, and its relationship with vital anatomical structures. While a conservative approach may be sufficient in cases with minimal symptoms, surgical intervention is required in cases of progressive pain, the presence of necrotic tissue, and spread into anatomical spaces.[8] In the present case, progressive pain, cortical perforation, extensive necrotic tissue formation, and extension into the submandibular space constituted clear indications for surgical intervention. Surgical debridement allows for the control of inflammation by removing extruded materials and eliminating necrotic tissues. This approach plays a critical role in preventing the progression of complications of the disease.[9] In the present case, a successful clinical outcome was achieved with early surgical treatment. This supports the decisive impact of early diagnosis and appropriate surgical approaches on prognosis. Furthermore, the current literature shows that delayed intervention in similar cases is associated with greater tissue loss.[10]

A limitation of the present report is the relatively short follow-up period, which precluded assessment of long-term bone remodeling and neurosensory outcomes.

CONCLUSION

Although calcium hydroxide is an effective and reliable intracanal medication for endodontic treatment, it can lead to serious iatrogenic complications in the event of uncontrolled extrusion. This case demonstrates that significant bone and soft tissue destruction can develop even when the inferior alveolar nerve is preserved. Particular attention should be paid to the proximity of anatomical structures in the mandibular posterior region, and excessive pressure should be avoided during medication administration. Early diagnosis, use of appropriate imaging methods, and determination of the correct treatment approach are critical for prognosis. In the presence of progressive symptoms and widespread tissue damage, successful clinical outcomes can be achieved through surgical debridement, removal of extruded material, and elimination of necrotic tissue. Therefore, clinicians must be aware that complications associated with calcium hydroxide extrusion are not limited to neurosensory findings and should be attentive to early intervention.

ETHICAL DECLARATIONS

Informed Consent

Informed consent was obtained from the legal guardians of the pediatric patient described in this report. Where developmentally appropriate, assent was also sought from the child. The inclusion of vulnerable populations in this study adhered to national and international ethical guidelines. Extra care was taken to ensure voluntary participation, understanding, and protection of participant dignity and autonomy.

Peer Review Process

This report underwent external peer review.

Conflict of Interest

The authors declare no conflicts of interest.

Financial Disclosure

This case report did not receive any financial support.

Author Contributions

Concept: HİD, MSŞ, TK, KFK; Design: HİD, MSŞ; Control: HİD, MSŞ; Data Collection and/or Processing: TK, KFK; Analysis and/or Interpretation: HİD, MSŞ; Literature Review: TK, KFK; Writing the Article: HİD, MSŞ, TK, KFK; Critical Review: HİD, MSŞ, TK, KFK.

Acknowledgments

The authors thank all individuals who contributed to this study.

REFERENCES

1. Siqueira JF Jr, Rôças IN. Clinical implications and microbiology of bacterial persistence after treatment. *J Endod.* 2008;34(11):1291-1301.e3. doi:10.1016/j.joen.2008.07.028
2. Mohammadi Z, Dummer PM. Properties and applications of calcium hydroxide in endodontics and dental trauma. *Int Endod J.* 2011;44(8):697-730. doi:10.1111/j.1365-2591.2011.01886.x
3. Montenegro Fonsêca J, Rangel Palmier N, Amaral-Silva GK, et al. Massive extrusion of calcium hydroxide paste containing barium sulfate during endodontic treatment. *Aust Endod J.* 2020;46(2):257-262. doi:10.1111/aej.12382
4. Gomes BP, Vianna ME, Sena NT, Zaia AA, Ferraz CC, de Souza Filho FJ. In vitro evaluation of the antimicrobial activity of calcium hydroxide combined with chlorhexidine gel as an intracanal medicament. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod.* 2006;102(4):544-550. doi:10.1016/j.tripleo.2006.04.010
5. Kasapoğlu MB, Doğancalı GE. Inferior alveolar nerve injury due to extrusion of calcium hydroxide during endodontic treatment: a case report. *Aust Endod J.* 2022;48(2):342-346. doi:10.1111/aej.12650
6. Almaslamani M, Alhassawi S, Prasad P, Alshayeb M. Diagnosis and management of inferior alveolar nerve injury caused by calcium hydroxide extrusion during root canal treatment: case report. *J Pharm Bioallied Sci.* 2024;16(Suppl 5):S4900-S4904. doi:10.4103/jpbs.jpbs_160_24
7. Patel S, Dawood A, Ford TP, Whaites E. Potential applications of cone beam computed tomography in the management of endodontic problems. *Int Endod J.* 2007;40(10):818-830. doi:10.1111/j.1365-2591.2007.01299.x
8. Deliverska E, Pashova-Tasseva Z. Sealer extrusion and osteomyelitis. *Pharmacia.* 2026;73:1-8. doi:10.3897/pharmacia.73.e185175
9. Hegde VR, Jain A, Patekar SB. Comparative evaluation of calcium hydroxide and other intracanal medicaments on postoperative pain in patients undergoing endodontic treatment: a systematic review and meta-analysis. *J Conserv Dent.* 2023;26(2):134-142. doi:10.4103/jcd.jcd_501_22
10. Tripathi S, Tripathi P, Misra P, Divyanu. Accidental extrusion of calcium hydroxide into the inferior alveolar canal during endodontic treatment: a case report. *J Pharm Bioallied Sci.* 2025;17(Suppl 2):S1984-S1987. doi:10.4103/jpbs.jpbs_1386_24