RESEARCH ARTICLE

Recevied/ Geliş Tarihi: 18.06.2025
Accepted/ Kabul Tarihi: 30.07.2025
Published / Yayın Tarihi: 31.08.2025

Aydın Dental Journal - Volume **11** Issue **2** - August **2025** (139-146)

Aydın Dental Journal

Journal homepage: http://dergipark.ulakbim.gov.tr/adj
DOI: 10.17932/IAU.DENTAL.2015.009/dental_v011i2003

Dentistry Students vs. ChatGPT 4o: Assessing Knowledge of Radiographic Anatomy on Panoramic Images

Diş Hekimliği Öğrencileri ve ChatGPT 4o: Panoramik Görüntülerde Radyografik Anatomi Bilgisinin Değerlendirilmesi

Melisa Öçbe[1*], Yeşim Demirdurak[2]

ABSTRACT

Objectives: Panoramic radiography constitutes an essential modality for the comprehensive assessment of patients in dentistry. This study investigates the panoramic radiography knowledge between dentistry students and ChatGPT-4o.

Materials and Methods: In this study, 182 undergraduate students from Kocaeli Health and Technology University, Faculty of Dentistry participated, with 52.7% being 3rd grade students and 47.3% being 4th grade students. Participants identified 35 landmarks on panoramic radiographs, which were categorized into bone structures, soft tissues/airways, and ghost images (20, 13, 2 questions were asked respectively). ChatGPT-4o was evaluated on the same task using the identical set of panoramic radiographic anatomy landmarks. To ensure consistency, the questions were entered into the AI model in the same format as they were presented to the students. ChatGPT-4o's responses were recorded for further evaluation. All data were analyzed using IBM SPSS version 23.

Results: The study revealed that 4th-grade students demonstrated significantly higher correct answer rates in bone structures, ghost images, and total scores compared to 3rd-grade students. No significant differences were observed in the median values of correct answer rates according to age and gender.

Conclusions: The outcome that 4th-grade students exhibited greater competence compared to 3rd-graders suggests that practical exposure to panoramic radiography within clinical contexts may enhance knowledge retention despite comprehensive theoretical instruction.

Keywords: *Anatomy, Artificial intelligence, Dental education, Panoramic radiography.*

ÖZET

Amaç: Panoramik radyografi, diş hekimliğinde hastaların kapsamlı değerlendirilmesinde temel bir görüntüleme yöntemidir. Bu çalışmada, diş hekimliği öğrencileri ile 4o'nun panoramik radyografi bilgileri karşılaştırılmıştır.

Gereç ve Yöntemler: Bu çalışmaya Kocaeli Sağlık ve Teknoloji Üniversitesi Diş Hekimliği Fakültesi'nden toplam 182 lisans öğrencisi katılmış olup, katılımcıların %52,7'si 3. sınıf, %47,3'ü ise 4. sınıf öğrencilerinden oluşmaktadır. Katılımcılardan panoramik radyografiler üzerinde 35 anatomik oluşumu tanımlamaları istenmiş ve bu yapılar; kemik yapılar, yumuşak dokular / hava yolları ve hayalet görüntüler olarak kategorize edilmiştir (sırasıyla 20, 13 ve 2 soru sorulmuştur). ChatGPT-4o, aynı panoramik radyografik anatomi yapıları kullanılarak aynı görevde değerlendirilmiştir. Tutarlılığı sağlamak amacıyla sorular, öğrencilere sunulduğu formatta yapay zeka modeline girilmiştir. ChatGPT-4o'nun yanıtları kaydedilmiş ve daha sonra analiz edilmiştir. Tüm veriler IBM SPSS 23 kullanılarak analiz edilmiştir.

Bulgular: Çalışmada, 4. sınıf öğrencilerinin kemik yapılar, hayalet görüntüler ve toplam puanlarda 3. sınıf öğrencilerine kıyasla anlamlı düzeyde daha yüksek doğru cevap oranlarına sahip olduğu bulunmuştur. Yaş ve cinsiyete göre doğru cevap oranlarının medyan değerlerinde anlamlı bir fark saptanmamıştır.
Sonuç: Dördüncü sınıf öğrencilerinin üçüncü sınıf öğrencilerine kıyasla daha yüksek başarı göstermesi, panoramik radyografinin klinik uygulamalarda kullanılmasıyla sağlanan pratik deneyimin, kapsamlı teorik eğitime rağmen bilgi kalıcılığını artırabileceğini düşündürmektedir.

Anahtar Kelimeler: *Anatomi, Diş hekimliği eğitimi, Panoramik radyografi, Yapay zeka.*

[1*]Assistant Professor, Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Turkiye.
[2] Assistant Professor, Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Turkiye.

***Corresponding Author:** Melisa Ocbe**, e-mail:** melisa.ocbe@kocaelisaglik.edu.tr, **ORCID:** 0000-0003-1609-610X, Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Turkey.

Introduction

Dental education in Türkiye is a 5-year program, and the comprehensive curriculum includes foundational lectures and basic training throughout the five years, with clinical preliminary training initiating in the third year, and clinical training in the fourth and fifth years, with intraoral and extraoral radiographic anatomy lectures during their third year. These courses provide students with essential knowledge to understand dentomaxillofacial anatomical structures. When studying these anatomical structures, students often find certain regions challenging to understand and remember which can hinder their ability to accurately identify these areas in clinical settings. Although previous studies have clearly shown that dental students experience notable difficulty in identifying certain anatomical regions on panoramic radiographs, especially soft tissue structures, ghost images, and projection artifacts,[1,2] there are still only a limited number of studies that systematically evaluate these challenges by anatomical region. Given the inherent complexity of panoramic imaging and its reliance on superimposed two-dimensional representations of three-dimensional structures, a more detailed and region-specific assessment is essential to improve radiographic anatomy education in dental curricula.

Chat Generative Pretrained Transformer (ChatGPT) is an artificial intelligence model capable of not only generating text but also analyzing visuals and interpreting medical images.[3,4] While numerous studies in the literature explore ChatGPT's applications in medical image analysis [5-8], its use in panoramic radiography remains limited. The aim of this study is to compare the knowledge of panoramic radiographic anatomy between dentistry students and the current version of ChatGPT, 4o.

Materials and Methods

This study involved 182 undergraduate students (87% of all invited participants) from the Kocaeli Health and Technology University Faculty of Dentistry, targeting the 3rd and 4th graders. Data were collected through voluntary participation in the study. Participants were informed about the purpose of the study and provided consent for their inclusion. Informed consent was obtained from all subjects involved in the study. Students who did not provide informed consent, those who were not enrolled in the 3rd or 4th year of the dentistry program at the time of the study, and individuals who did not complete the full set of questions or left the assessment incomplete were excluded from the study. Ethical approval for this study was obtained from the Institutional Ethical Committee of Kocaeli Health and Technology University (Project Number: 2024/79). All procedures were conducted in accordance with the principles outlined in the Declaration of Helsinki.

The anatomical landmarks were categorized into three main groups: bone structures, soft tissue/airways, and ghost images. In total, 35 anatomical landmarks were evaluated-comprising 20 bone structures, 13 soft tissue or airway structures, and 2 ghost images. Importantly, only one panoramic radiograph was used throughout the study. Each of the 35 anatomical points was marked individually on this single image, and participants were asked to identify the anatomical structure corresponding to each marked point, one by one.

As illustrated in Figure 1, the same panoramic radiograph was used for all questions, with a different anatomical point highlighted for each question. To annotate the 35 distinct anatomical landmarks on a single panoramic radiograph, the Segment Editor module of the 3D Slicer software (version 5.2.2, USA) was used. During the face-to-face sessions, the panoramic image was projected on a screen, and participants were shown each marked point in succession. For every question, students were asked, “Which anatomical structure is indicated by the marked point?” This process ensured that students interpreted the same image throughout, but focused on a different anatomical landmark each time.

All assessments were conducted under the supervision of two observers-an Oral and Maxillofacial Radiologist with 6 years of experience (MO) and another with 27 years of experience (YD)-to ensure consistency and reliability. Correct and incorrect identifications were recorded for each landmark.

To evaluate AI performance, the same 35 annotated questions-based on the single panoramic image-were presented to ChatGPT-4o. The model received the questions in the exact same format used for students to maintain methodological consistency.

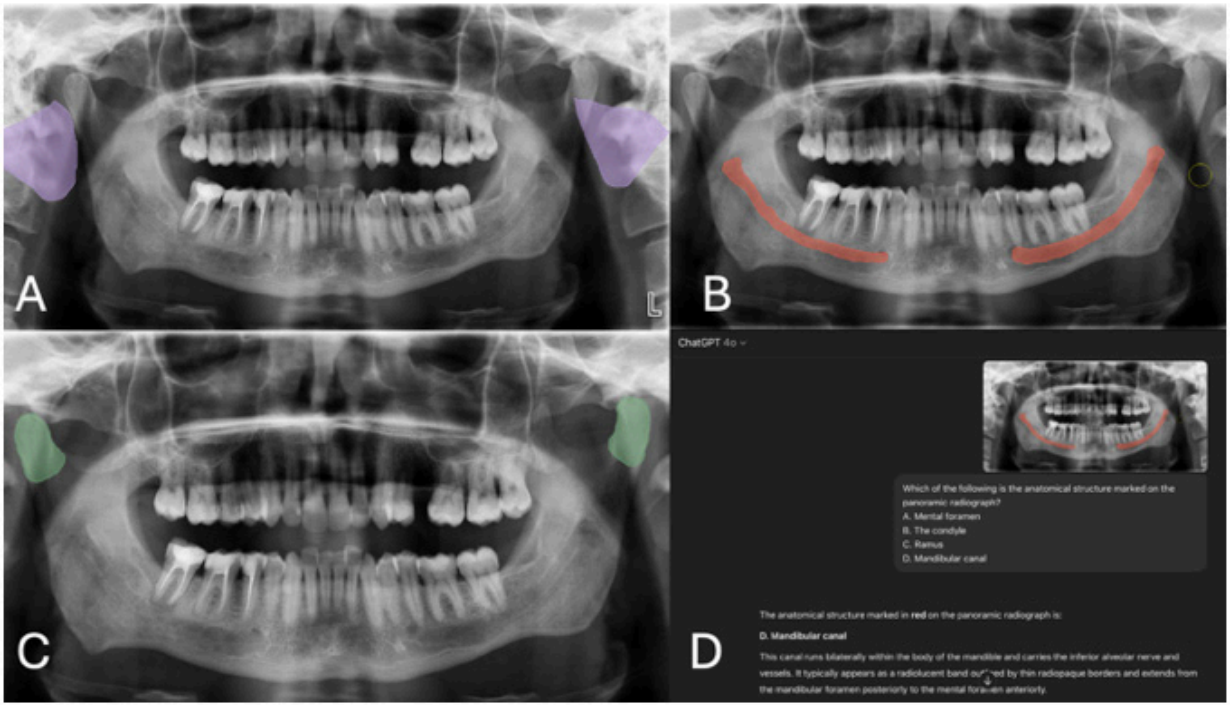

Figure 1. A. Panoramic radiograph showing annotated earlobes. B. Highlighted mandibular canal bilaterally. C. Bilateral condyles visualized at the superior aspect of the mandible. D. Screenshot of the question format presented to ChatGPT and the model's correct identification of the mandibular canal.

The data were analysed by IBM SPSS V23. The conformity of the data to normal distribution was analysed by Kolmogorov Smirnov test. Mann Whitney U test was used to compare the data that did not show normal distribution according to binary groups. Spearman's rho correlation coefficient was used to analyse the relationship between the scores that did not conform to normal distribution and age. Descriptive statistics results were presented as frequency (percentage) for categorical variables, mean ± standard deviation and median (minimum - maximum) for quantitative variables. Significance level was taken as $p<0.050$.

Results

The participants consisted of 98 female students (53.8%) and 82 male students (45.1%). Detailed descriptive statistics results are provided in Table 1. Statistically, no significant difference was observed in the median values of correct answer rates of hard tissue structures, soft tissue / airways, ghost images and total scores according to gender ($p=0.865$, $p=0.682$, $p=0.839$, $p=0.985$, respectively).

Table 1: Demographic characteristics and knowledge levels of dental students regarding panoramic radiographic anatomy.

		Frequency	%
Gender			
Female		98	53,8
Male		82	45,1
Non declared		2	1,1
Grade			
3rd Grade		96	52,7
4th Grade		86	47,3
		Mean ± Standard Deviation	**Median (Min-Max)**
Age		22,92 ± 3,28	22 (19 - 39)
Correct answer rates			
	Bone sturcture	18,13 ± 2,32	19 (7 - 20)
	Ghost image	1,83 ± 0,42	2 (0 - 2)

	Soft tissue / Airways	11,36 ± 1,77	12 (5 - 13)
	Total score	31,32 ± 3,79	33 (13 - 35)

The median number of correct answers for bone structure questions among 3rd-grade students was slightly lower compared to 4th-grade students. The median values exhibited significant differences based on the grade factor (p<0.001). The median number of correct answers for ghost image questions was 2 for both 3rd and 4th-grade students. The distinction arised from the average score, with 3rd-grade students achieving a rank average value of 96.18 compared to 86.28 for 4th-grade students (p=0.043). The median number of correct answers for soft tissue and airways questions was consistent across both 3rd and 4th-grade students, both recording a median value of 12 (p=0.315). Regarding the total score, the median value was 31 for 3rd-grade students and 33 for 4th-grade students, demonstrating a significant difference between the two groups (p=0.008) (Table 2). When examined through Spearman correlation, there was no statistically significant relationship observed between age and correct answer rates of bone structures (r=0.058; p=0.439), ghost images (r=-0.049; p=0.514), soft tissue / airways (r=-0.037; p=0.616), and total score (r=0.041; p=0.585).

Table 2: Correct answer rates of panoramic radiographic anatomy by grade level

Correct answer rates of:	3rd Grade	4th Grade	Total	Test StatIstIcs	p
Bone structures	18 (7 - 20)	19 (11 - 20)	19 (7 - 20)	2928,000	<0,001*
Ghost Images	2 (1 - 2)	2 (0 - 2)	2 (0 - 2)	4577,000	0,043*
Soft tissue / AIrways	12 (5 - 13)	12 (6 - 13)	12 (5 - 13)	3782,500	0,315*
Total score	31 (13 - 35)	33 (19 - 35)	33 (13 - 35)	3199,000	0,008*

**Mann Whitney U Test*

The pterygopalatine fossa, inferior nasal meatus, and middle nasal meatus emerged as the least recognized anatomical landmarks, with correct answer rates of 64.8%, 65.4%, and 73.1%, respectively. Table 3 and Table 4 demonstrates detailed information on correct answer rates, most common wrong answer, and rate of the most common wrong answer and the answers of Chat GPT 4o. Chat GPT 4o correctly answered 25 of 35 questions (71%). Inferior nasal concha and inferior nasal meatus were most common mistakes of both dentistry students and Chat GPT 4o.

Table 3: Correct identification rates and common errors for anatomical landmarks of bone structures.

	Anatomical landmark	**Rate of the correct answer (%)**	**Most common wrong answer of dentistry students**	**Rate of the most common wrong answer (%)**	**Answer of Chat GPT 4O**
Bone structures	Condyle	96.7	Coronoid process	2.7	Correct
	Coronoid process	96.7	Maxillary sinus	3.3	***Maxillary sinus***
	Ramus	97.8	Mandibular canal	1.6	Correct
	Inferior border of the mandible	96.2	Styloid process	3.8	Correct
	Articular eminence	86.8	Glenoid fossa	11	***Glenoid fossa***
	External acusticus meatus	83	Styloid process	8.8	Correct
	Mandibular foramen	94.5	Infraorbital canal	3.8	***Mental foramen***
	Mental foramen	95.1	Mandibular foramen	3.3	Correct
	Mandibular canal	98.9	Mental foramen	1	Correct
	Hyoid bone	96.2	Angle of the mandible	3.8	Correct

	Styloid process	82.4	External acusticus meatus	14.8	Correct
	Infraorbital canal	83.5	Pterygopalatine fossa	8.2	***Floor of the maxillary sinus***
	Floor of orbita	88.5	Floor of maxillary sinus	11	***Floor of the maxillary sinus***
	Pterygopalatine fossa	64.8	Ethmoid sinus	17	***Maxillary Sinus***
	Sigmoid notch	90.7	Coronid process	5.5	***Condyle***
	Zygoma	90.7	Maxillary sinus	6	Correct
	Anterior nasal spine	98.9	Maxillary sinus	1	Correct
	Pterygopalatine fissure	94	Inferior nasal concha	4.4	Correct
	Hard palate	92.9	Soft palate	3.8	Correct
	Glenoid fossa	85.2	Condyle	8.8	Correct

Table 4: Correct identification rates and common errors for soft tissue anatomical landmarks and ghost images.

	AnatomIcal landmark	**Rate of the correct answer (%)**	**Most common wrong answer**	**Rate of the most common wrong answer (%)**	**Answer of Chat GPT 4O**
Soft tIssues and aIrways	Dorsum of the tongue	78	Soft palate	17	Correct
	Soft palate	87.9	Hard palate	7.1	Correct
	Nasal septum	98.9	Nasopharynx	1	Correct
	The earlobes	94.5	Hypopharynx	5.5	Correct
	The nose	96.7	Soft palate	2.2	Correct
	The posterior wall of the oral and nasal pharynx	96.2	Inferior nasal meatus	2.2	Correct
	Middle nasal concha	98.4	Mental foramen	1.6	Correct
	Middle nasal meatus	73.1	Anterior nasal spine	18.7	***Maxillary sinus***
	Inferior nasal concha	91.2	Orbita	6.6	***Orbita***
	Inferior nasal meatus	65.4	Floor of the maxillary sinus	22.5	***Floor of the maxillary sinus***
	Hypopharynx	95.6	Mandibular canal	3.3	Correct
	Nasopharynx	83.5	Styloid process	13.2	Correct
	Oropharynx	76.4	Mastoid air cells	12.1	Correct
	Ghost image of the cervical vertebrae	95.1	Hypopharynx	2.7	Correct
	Ghost image of contralateral mandible	87.9	Corpus mandible	11	Correct

Discussion

In this study, it was found that 4th-year students who had begun clinical education demonstrated a higher level of knowledge compared to 3rd-year students who had completed radiographic anatomy courses but had not yet started clinical education. In addition, Chat GPT 4o represented a 71% of correct answers, which was found to be lower than both 3rd and 4th grade dentistry students. It was observed that the anatomical structures with the lowest recognition rates for dentistry students were the pterygopalatine fossa, inferior nasal meatus, and middle nasal meatus. Knowledge of hard tissues was generally more prevalent in dentistry students, whereas the identification and understanding of soft tissues and fossae were less well-acquired by dental students.

Although, Chat GPT 4o scored higher in soft tissue and ghost image questions.

The study revealed specific areas where students and Chat GPT 4o demonstrated common misunderstandings. For instance, 96.7% of the students correctly identified the mandibular condyle, yet there was a notable confusion with the coronoid process among those who answered incorrectly. Chat GPT 4o answered mandibular condyle question correctly; while did not answer correctly for coronoid process. Similarly, while 86.8% of participants accurately identified the articular eminence, the glenoid fossa was frequently mistaken by those who erred. Chat GPT 4o also misidentified articular eminence as glenoid fossa. These results highlight a tendency for students to confuse anatomical structures that are in close proximity to each other on radiographs, with a need for increased focus on teaching and reinforcing the identification these structures in dental radiology education.

In a previous research by Serindere and Aktuna[9], it was stated that mandibular canal, mandibular foramen, mental foramen, and incisive canal are the most visible anatomical structures on panoramic radiographs.[9] In addition, the findings of Tokgoz Kaplan and Yıldız[10] indicated that the mental foramen had the highest correct identification rate among dental students.[10] Our study also revealed high levels of knowledge regarding these structures, confirming their visibility and recognizability in radiographic images. However, despite their distinct appearances, the mental foramen and mandibular canal were often mistaken for one another by the students. This confusion highlights a persistent issue in dental radiographic education where students struggle to differentiate between similarly located structures, despite their differing radiographic presentations. Moreover, our findings resonate with the studies of Maeda et al. and Elsheikh et al.[11], which identified the nasal cavity structures as challenging for dental students to detect on panoramic radiographs.[1,11] Specifically, our study found that knowledge levels were particularly low for the middle and inferior nasal meatus. This suggests a need for improved educational strategies focusing on the nasal cavity structures to enhance the diagnostic accuracy of dental students. In light of these findings, it is evident that a targeted educational focus on differentiating between closely situated anatomical landmarks is necessary. Enhanced training methods, such as side-by-side comparisons, virtual reality simulations, artificial intelligence-assisted softwares and detailed annotations on radiographs, could be beneficial.[12-14] By addressing these specific areas of confusion, dental education can improve the accuracy of anatomical identification and ultimately enhance the diagnostic capabilities of future dental professionals.

Conclusions

This study evaluated and compared the knowledge of panoramic radiographic anatomy among third- and fourth-year dental students and the ChatGPT 4o large language model. The findings demonstrated that fourth-year students, who had begun their clinical education, showed significantly better performance compared to third-year students, indicating that clinical exposure enhances radiographic anatomical understanding. ChatGPT 4o achieved a 71% overall accuracy rate, which was lower than the performance of both student groups, suggesting that while large language models (LLMs) show promise, their current capabilities do not yet surpass those of trained dental students in this specific task.

The results further revealed specific anatomical regions-particularly the pterygopalatine fossa and nasal meatus areas-as common sources of misidentification, highlighting a gap in students' understanding of complex or less prominent anatomical structures. Similar trends in errors were observed in both students and the AI model, particularly in distinguishing between anatomically adjacent features such as the mandibular condyle and coronoid process, or the articular eminence and glenoid fossa. Incorporating advanced tools such as AI-assisted visualizations, side-by-side radiographic comparisons, and interactive learning technologies may help address these gaps. Although ChatGPT 4o exhibited potential in anatomical interpretation, its limitations reinforce the importance of human expertise and structured anatomical training in dental education. Future research may explore the integration of AI as a supportive tool in dental curricula and assess its impact on student learning outcomes.

Acknowledgement

None declared.

Ethical Approval

Ethical approval for this study was obtained from the Institutional Ethical Committee of Kocaeli Health and Technology University (Project Number: 2024/79) at 17.01.2024.

Funding

The authors declare that this research received no

specific grant from any funding agency in the public, commercial, or not-for-profit sectors and that they have no financial interests related to the study.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding this study.

Authorship Contributions

Idea/Concept: M.O, Y.D Design: M.O, Y.D Control/ Supervision: M.O, Y.D Literature Review: Y.D Data Collection and/or Processing: M.O, Y.D Analysis and/or Interpretation: Y.D Writing the Article: M.O Critical Review: M.O

References

1. Maeda N, Hosoki H, Yoshida M, Suito H, Honda E. Dental students' levels of understanding normal panoramic anatomy. *J Dent Sci* 2018; 13: 241-6.
2. Razmus TF, Williamson GF, Van Dis ML. Assessment of the knowledge of graduating American dental students about the panoramic image. *Oral Surg Oral Med Oral Pathol* 1993; 76: 499-504.
3. Tran CG, Chang J, Sherman SK, De Andrade JP. Performance of ChatGPT on American Board of Surgery In-Training Examination preparation questions. *J Surg Res* 2024; 291: 61-7.
4. Gordon EB, Towbin AJ, Wingrove P, et al. Enhancing patient communication with Chat-GPT in radiology: Evaluating the efficacy and readability of answers to common imaging-related questions. *J Am Coll Radiol* 2024; 21: 142-9.
5. Mago J, Sharma M. The potential usefulness of ChatGPT in oral and maxillofacial radiology. *Cureus* 2023; 15: e42133.
6. Khurana S, Vaddi A. ChatGPT from the perspective of an academic oral and maxillofacial radiologist. *Cureus* 2023; 15: e40053.
7. Baghdady MT, Pharoah MJ, Regehr G, Lam EW, Woods NN. The role of basic sciences in diagnostic oral radiology. *J Dent Educ* 2009; 73: 1186-93.
8. Parillo M, Vaccarino F, Beomonte Zobel B, Mallio CA. ChatGPT and radiology report: potential applications and limitations. *Radiol Med* 2024; 129: 527-33.
9. Serindere G, Belgin CA. The evaluation of visibility of mandibular anatomic landmarks using panoramic radiography. *Int Dent Res* 2019; 9: 60-5.
10. Tokgöz Kaplan T, Yıldız Y. Evaluation of knowledge levels of undergraduate and postgraduate dentists concerning anatomical landmarks on panoramic radiography. *Unika Sağlık Bilimleri Dergisi* 2024; 8: 671-81.
11. Elsheikhi F, Betamar N, Bugaighis I. Assessment of level of knowledge understood and memorized of panoramic anatomical landmarks among dental students in Benghazi, Libya. *Libyan Int Med Univ J* 2021; 6: 145-50.
12. Alsufyani N, Alnamlah S, Mutaieb S, et al. Virtual reality simulation of panoramic radiographic anatomy for dental students. *J Dent Educ* 2023; 87: 1208-14.
13. Wang D, Huai B, Ma X, et al. Application of artificial intelligence-assisted image diagnosis software based on volume data reconstruction technique in medical imaging practice teaching. *BMC Med Educ* 2024; 24: 145.
14. Yuan Q, Chen X, Zhai J, et al. Application of 3D modeling and fusion technology of medical image data in image teaching. *BMC Med Educ* 2021; 21: 316.