

DOI: <http://dx.doi.org/10.12996/gmj.2026.4729>

Study Habits and Attitudes Toward Artificial Intelligence Tools Among Anesthesia Trainees in Türkiye: A Descriptive Survey

Türkiye'deki Anestezi Asistanlarında Çalışma Alışkanlıkları ve Yapay Zeka Araçlarına Yönelik Tutumlar: Tanımlayıcı Bir Anket Çalışması

© Çağrı Özdemir¹, © Beyza Büyükgebiz Yeşil², © Miray Gözde Özdemir³

¹Department of Anesthesiology and Reanimation, Gazi University, Faculty of Medicine, Ankara, Türkiye

²Clinic of Anesthesiology and Reanimation, Kahramanmaraş State Hospital, Kahramanmaraş, Türkiye

³Department of Anesthesiology and Reanimation, University of Health Sciences Türkiye, Ankara Etlik City Hospital, Ankara, Türkiye

ABSTRACT

Objective: To describe the study habits and attitudes toward the use of artificial intelligence (AI) tools for studying among anesthesia trainees in Türkiye, and to explore the associations between these factors and performance in national in-training examinations.

Methods: A cross-sectional online survey was administered to anesthesia trainees in Türkiye between April and June 2025. The questionnaire addressed study frequency and duration, preferred resources, AI use patterns, and 5-point Likert ratings for reliability and usefulness. Performance in the national online trainee assessment examination (ÇADS) was analyzed using Kruskal-Wallis tests and Spearman correlation, with ÇADS scores encoded as ordinal ranks (1-4).

Results: A total of 148 responses were analyzed. Study sessions were reported as "occasional" by 60.1%, with typical sessions lasting 1-2 hours (48.6%). Although 96.6% of the respondents perceived a need to study more, the main barriers were fatigue and burnout (46.9%) and time constraints due to clinical duties (29.4%). The most commonly used resources were textbooks (92.6%), spaced repetition (60.8%), and online journals (52.7%). AI tools were rated as useful (mean 3.89 ± 0.99) but only moderately reliable (3.22 ± 0.97). Two-thirds of the respondents participated regularly in ÇADS (66.2%). Among the 130 participants who had completed at least one ÇADS examination, the most frequently reported score category was 60-79 (73.1%). Frequent AI use was significantly associated with higher ÇADS performance ($p =$

ÖZ

Amaç: Türkiye'de anesteziyoloji ve reanimasyon asistanlarının klinik dışı çalışma alışkanlıklarını, kullandıkları eğitim kaynaklarını ve yapay zeka (YZ) tabanlı araçlara ilişkin kullanım sıklığını ortaya koymak, ayrıca bu faktörlerin ulusal çevrim içi asistan değerlendirme sınavı (ÇADS) performansı ile ilişkisini değerlendirmektir.

Yöntemler: Nisan-Haziran 2025 tarihleri arasında anesteziyoloji asistanlarına çevrimiçi, anonim, kesitsel bir anket uygulandı. Anket; demografik özellikler, çalışma sıklığı ve süresi, eğitim kaynakları, YZ kullanım sıklığı ve 5'li Likert ölçeği ile güvenilirlik/fayda algısı sorularından oluştu. ÇADS skorları sıralı kategoriler (1-4) olarak kodlandı ve Kruskal-Wallis testi ve Spearman korelasyonu ile değerlendirildi.

Bulgular: Toplam 148 yanıt analiz edildi. Çalışma sıklığı ağırlıklı olarak "ara sıra" (%60,1), oturum süresi 1-2 saat (%48,6) olarak bildirildi. Katılımcıların %96,6'sı daha fazla çalışması gerektiğini düşünüyordu; ana engeller yorgunluk/tükenmişlik (%46,9) ve klinik görevlerden kaynaklı zaman eksikliği (%29,4) idi. En sık kullanılan kaynaklar ders kitapları (%92,6), aralıklı tekrar (%60,8) ve çevrimiçi makaleler (%52,7) iken YZ araçlarını düzenli kullananlar yalnızca %24,3 oranındaydı. YZ'nin faydalılığı ortalama $3,89 \pm 0,99$, güvenilirliği $3,22 \pm 0,97$ olarak puanlandı. ÇADS sınavına katılan 130 asistanın %73,1'inin son ÇADS puanı 60-79 aralığındaydı. Düzenli YZ kullanımı (günlük veya haftada 2-3 kez) ÇADS performansı ile anlamlı ilişkiliydi ($p = 0,003$). Kendine güven düzeyi de ÇADS performansı ile pozitif korele idi ($r_s = 0,205$, $p = 0,019$). Çalışma sıklığı ve kullanılan teknikler ÇADS performansı ile ilişkili değildi.

Cite this article as: Özdemir Ç, Büyükgebiz Yeşil B, Özdemir MG. Study habits and attitudes toward artificial intelligence tools among anesthesia trainees in Türkiye: a descriptive survey. Gazi Med J. 2026;37(4):496-503

Address for Correspondence/Yazışma Adresi: Çağrı Özdemir, Department of Anesthesiology and Reanimation, Gazi University, Faculty of Medicine, Ankara, Türkiye

E-mail / E-posta: cagrioedemir@gazi.edu.tr

ORCID ID: orcid.org/0000-0002-1266-8054

Received/Geliş Tarihi: 02.06.2026

Accepted/Kabul Tarihi: 17.08.2026

Publication Date/Yayınlanma Tarihi: 15.10.2026



©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Gazi University Faculty of Medicine. Licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

©Telif Hakkı 2026 Yazar(lar). Gazi Üniversitesi Tıp Fakültesi adına Galenos Yayınevi tarafından yayımlanmaktadır. Creative Commons Atıf-GayriTicari-Türetilemez 4.0 (CC BY-NC-ND) Uluslararası Lisansı ile lisanslanmaktadır.

ABSTRACT

0.003). No associations were found between ÇADS performance and study frequency or specific study techniques.

Conclusion: Anesthesia trainees in Türkiye display irregular, brief study patterns, with fatigue and burnout as the main problems. AI tool adoption remains low, likely due to moderate trust and a lack of guidance. Frequent AI use and higher self-confidence were associated with better examination performance in exploratory univariable analyses. Educational programs may benefit from addressing workload and burnout while providing structured guidance for the appropriate integration of AI tools.

Keywords: Anesthesia trainees, study habits, artificial intelligence, e-learning, Türkiye

Öz

Sonuç: Türkiye'deki anesteziyoloji asistanları düzensiz ve kısa süreli çalışma örüntüsü sergilemekte; yorgunluk ve tükenmişlik temel sorunlar olarak öne çıkmaktadır. YZ araçları düşük güven ve rehberlik eksikliği nedeniyle az kullanılmaktadır. Tek değişkenli keşifsel analizlerde, sık YZ kullanımı ve yüksek kendine güven düzeyinin daha iyi sınav performansı ile ilişkili olduğu bulundu. Eğitim programları iş yükü ve tükenmişliği doğrudan ele almalı, YZ araçlarının entegrasyonu desteklenmelidir.

Anahtar Sözcükler: Anesteziyoloji, çalışma alışkanlıkları, yapay zeka, çevrimiçi öğrenme, Türkiye

INTRODUCTION

In medical education, trainee study behavior is not uniform across specialties, institutions, or even training years; it is an adaptive response to workload, assessment pressure, and the availability of structured learning tools.

Beyond bedside learning, trainees are expected to engage in self-directed study to maintain clinical competence and prepare for in-training or board examinations. Accordingly, study habits constitute an important component of residency education, as they shape not only examination preparation but also the development of long-term learning routines (1-4).

Previous studies from different specialties suggest that trainees' study behaviors are influenced by multiple factors, including workload, duty-hour structure, educational resources, and the assessment system within their training program (1-5). Trainees often rely on a combination of traditional methods, such as textbooks and written summaries, and more structured or exam-oriented approaches, including case conferences, question-based learning, and online modules (1-3). Access to digital resources and structured educational formats has also been associated with more consistent study patterns and, in some reports, with improved examination performance (1,2,5). At the same time, external pressures such as heavy service demands, fatigue, and disrupted clinical environments may reduce the time and energy available for self-study (5-7).

These issues may be particularly relevant in anesthesiology, a specialty characterized by intense clinical responsibility, irregular working hours, night shifts, and a high cognitive load. Although anesthesiology trainees are expected to master broad theoretical knowledge while simultaneously developing procedural and decision-making skills, little is known about how they actually study outside clinical hours, which learning resources they prefer, and what barriers limit their study routines. This gap is important because educational planning is often based on assumptions rather than on direct data from the trainees themselves.

Since their public release in November 2022, artificial intelligence (AI) chatbots (also known as large language models or LLMs) have attracted substantial interest in medical education as fast and user-friendly tools that may support students, trainees, educators, and health care professionals in accessing and processing information (8-10). Early reports have shown that LLMs can achieve passing or

near-passing performance on components of medical examinations, including the United States Medical Licensing Examination, highlighting their potential relevance for knowledge acquisition and assessment (11,12). In addition to retrieving and summarizing information, LLMs may assist learners by generating explanations, clarifying difficult concepts, and supporting question-based learning in an interactive and low-stakes environment (9,10,12). Such tools may also facilitate academic tasks by providing rapidly synthesized information that can be used for studying and writing purposes (13). However, despite increasing interest in AI-assisted learning, the actual role of these tools in postgraduate medical education and their integration into trainees' everyday study routines remain insufficiently defined.

Despite growing interest in both trainee study behavior and digital learning tools, data from Türkiye remain limited. Findings from other countries may not be directly generalizable because residency training conditions vary substantially across health-care systems in terms of workload, institutional structure, educational culture, and access to learning resources. Understanding how anesthesiology trainees in Türkiye study, which resources they use, and how they perceive AI-based tools may provide useful insights for training programs seeking to improve learning support under real-world clinical constraints.

The present study aimed to evaluate the study habits of anesthesiology trainees in Türkiye, identify the main barriers affecting self-directed learning, describe the educational resources they use, and assess their attitudes toward the use of AI tools for studying and examination preparation.

MATERIALS AND METHODS

This was a national, cross-sectional, web-based survey conducted to explore the habits, preferred learning resources, and attitudes toward AI tools of anesthesiology trainees in Türkiye. Ethical approval was obtained from the Gazi University Ethics Commission (approval no: 13.01.2025-E.1141829, date: 13.01.2025). Participation was voluntary and anonymous, and all respondents provided electronic informed consent before proceeding to the survey. No financial or material incentives were offered. The study was reported in accordance with the Checklist for Reporting Results of Internet E-Surveys (14).

The questionnaire was developed by the authors specifically for this study to examine how anesthesiology trainees study outside clinical working hours and how they perceive emerging AI-based educational tools. After reviewing the relevant literature on resident study habits, examination preparation, digital learning resources, and AI-assisted education, the initial item pool was reviewed internally by the study team for relevance, clarity, and content coverage before pilot testing. Before distribution, the survey was pilot-tested among a small group of anesthesiology trainees who were not included in the final analysis, and minor wording revisions were made to improve clarity and completion flow. Although the questionnaire underwent internal review and pilot testing, it was not formally validated using psychometric methods; therefore, measurement error cannot be excluded. The final survey included items addressing demographic and training-related characteristics, frequency and duration of study outside clinical hours, commonly used study methods and educational resources, and patterns of AI tool use for learning and examination preparation. Additional questions addressed participation in national and international examinations, including the Turkish Society of Anaesthesiology and Reanimation online trainee assessment examinations (ÇADS) and the European Diploma in Anaesthesiology and Intensive Care (EDAIC).

The survey was prepared and administered through Typeform® (Typeform S.L., Barcelona, Spain), and the survey link was circulated electronically through WhatsApp groups used by anesthesiology trainees and other professional communication networks targeting anesthesiology trainees in Türkiye. This study therefore used a convenience sample of anesthesiology trainees who voluntarily responded to the survey through these professional communication networks. Because the exact number of trainees who received or viewed the invitation could not be determined, a formal response rate could not be calculated. Data collection was carried out between April and June 2025. Only fully completed questionnaires were included in the analysis, whereas incomplete submissions were excluded. Because the survey was conducted anonymously, no directly identifying personal information was retained in the final dataset.

No formal a priori sample size calculation was performed because the study was designed as an exploratory, web-based descriptive survey. The final sample size was comparable to those reported in previous survey-based studies evaluating study habits and academic performance among medical residents. For example, Barragán-Hervella et al. (4) analyzed 112 residents using a Likert-scale questionnaire and correlation-based analyses.

Statistical Analysis

Data were analyzed using StatsDirect (version 4.0.4, StatsDirect Ltd., UK). Categorical variables were summarized as numbers and percentages. Continuous variables were presented as mean $\pm$ standard deviation (SD), whereas Likert-scale items (1-5) were summarized using both mean $\pm$ SD and median with interquartile range (IQR). For analytical purposes, ÇADS score categories were encoded as ordinal ranks (1 = < 40, 2 = 40-59, 3 = 60-79, 4 = 80-100) and presented as medians with IQRs.

For group comparisons, the Kruskal-Wallis test was used for ordinal outcomes with more than two groups, followed by Dunn's

post-hoc test with Bonferroni correction where appropriate. The Mann-Whitney U test was applied for pairwise comparisons of the study techniques. Spearman's rank correlation coefficient (r_s) was calculated to assess associations between ordinal variables (self-confidence level and ÇADS performance). Categorical associations were examined using the chi-square test or Fisher's exact test where appropriate.

A two-sided p-value below 0.05 was considered statistically significant. All analyses were descriptive and exploratory in nature given the cross-sectional design.

RESULTS

A total of 148 responses were received and included in the analysis. The majority of respondents were aged 25-30 years ($n = 111$, 75.0%), with female participants comprising 56.8% ($n = 84$) of the sample. Detailed characteristics of the participants are shown in Table 1.

The majority of respondents reported studying outside clinical hours only "occasionally" ($n = 89$, 60.1%), while regular study (2-3 times per week or more frequently) was reported by less than one-quarter of the cohort. When study did occur, the typical session duration was 1-2 hours ($n = 72$, 48.6%) or less than one hour ($n = 57$, 38.5%), with sessions lasting longer than two hours reported by only 12.8% of participants. Despite infrequent and brief study patterns, almost all respondents ($n = 143$, 96.6%) reported that they believed they needed to study more. The most frequently cited barrier was fatigue and burnout ($n = 67$, 46.9%), followed by time constraints due to clinical duties ($n = 42$, 29.4%), lack of motivation ($n = 20$, 14.0%), and personal responsibilities outside work ($n = 14$, 9.8%) (Table 1).

Regular use of AI tools for study purposes (2-3 times per week or more) was reported by 24.3% of participants (Table 2). Among those who did use AI tools, ChatGPT or similar LLMs were by far the most commonly accessed ($n = 86$, 58.1%), while the use of clinical decision-support systems, AI-powered medical databases, or medical image analysis tools was uncommon ($\leq 8.1\%$). The most frequent purposes for AI use were information retrieval ($n = 84$, 56.8%), academic writing or presentation preparation ($n = 41$, 27.7%), and support with practice questions or case studies ($n = 27$, 18.2%).

The participants' perceptions of AI tools were characterized by moderate reliability [mean 3.22 ± 0.97 , median 3.0 (3.0-4.0)], higher perceived usefulness [mean 3.89 ± 0.99 , median 4.0 (3.0-5.0)], and perceived time-saving efficiency [mean 4.01 ± 1.01 , median 4.0 (3.0-5.0)]. Despite this, 67.6% of respondents rated AI tools as useful or very useful (scores of 4 or 5), yet actual usage remained limited, with 37.8% reporting no use at all (Table 2).

Textbooks were the most frequently reported learning resource ($n = 137$, 92.6%), followed by online articles or journals ($n = 78$, 52.7%) and educational videos ($n = 63$, 42.6%) (Table 3).

Among the 130 participants who had completed at least one ÇADS examination, the most frequently reported score category was 60-79 (95/130, 73.1%), while 26 participants (20.0%) reported scores between 80 and 100. Almost all participants ($n = 142$, 95.9%) had not yet attempted the EDAIC examinations (Table 3).

Exploratory analyses examined the relationship between study-related factors and ÇADS performance among the 130 participants

Table 1. Demographic and general characteristics of participants (n = 148).

Variable	n	%
Age group		
25-30	111	75.0
31-35	32	21.6
36-40	5	3.4
Sex		
Female	84	56.8
Male	64	43.2
Year of residency		
1 st year	26	17.6
2 nd year	55	37.2
3 rd year	22	14.9
4 th year	22	14.9
5 th year	23	15.5
Type of institution		
City hospital	93	62.8
University hospital	37	25.0
Training & research hospital	18	12.2
Study frequency (outside clinical hours)		
Every day	5	3.4
2-3 times/week	28	18.9
Once a week	26	17.6
Occasionally	89	60.1
Average study session duration		
<1 hour	57	38.5
1-2 hours	72	48.6
2-3 hours	16	10.8
>3 hours	3	2.0
Perceives need for more studying		
Yes	143	96.6
No	5	3.4
Main barrier to increased study (n = 143)		
Fatigue and burnout	67	46.9
Time constraints due to clinical duties	42	29.4
Lack of motivation	20	14.0
Personal responsibilities outside work	14	9.8

Data are presented as number (n) and percentage (%). Year of residency reflects the self-reported training year at the time of the survey.

who had completed at least one ÇADS examination (Table 4). ÇADS score categories were treated as ordinal ranks (1-4) for statistical comparison. AI tool use frequency was significantly associated with higher ÇADS scores (Kruskal-Wallis H = 15.76, p = 0.003). Post-hoc pairwise comparisons revealed that daily users [median score 3 (IQR: 3-4); 33.3% scoring ≥ 80] and users reporting 2-3 times per week (median 3 [IQR: 3-4]; 42.1% scoring ≥ 80) performed significantly

Table 2. Artificial intelligence tool use among anesthesiology trainees (n = 148).

Variable	n	%
Frequency of AI tool use for studying		
Every day	14	9.5
2-3 times/week	22	14.9
Once a week	8	5.4
Occasionally	48	32.4
Never	56	37.8
AI tools used (multiple response)		
ChatGPT or similar conversational AI	86	58.1
Clinical decision-support systems	1	0.7
AI-powered medical databases	12	8.1
Presentation-generating AI tools	6	4.1
Medical image analysis tools	8	5.4
Purposes of AI use (multiple response)		
Information retrieval	84	56.8
Practice questions/case studies	27	18.2
Academic writing/presentation	41	27.7
Exam preparation (EDAIC/ÇADS)	12	8.1
Simulation/guideline learning	22	14.9
Visual materials	25	16.9
AI evaluation scale (1-5)	Mean ± SD	Median [IQR]
Perceived reliability	3.22 (± 0.97)	3.0 [3.0-4.0]
Perceived usefulness	3.89 (± 0.99)	4.0 [3.0-5.0]
Time-saving efficiency	4.01 (± 1.01)	4.0 [3.0-5.0]

Data for frequency and tools are presented as n (%) or mean ± SD and median [IQR] as appropriate. Multiple-response items sum to more than 100%. AI evaluation scales: 1 = strongly disagree, 5 = strongly agree.

SD: Standard deviation, IQR: Interquartile range, AI: Artificial intelligence, EDAIC: European Diploma in Anaesthesiology and Intensive Care, ÇADS: Çevrimiçi Asistan Değerlendirme Sınavı.

better than non-users (median 3 [IQR: 3-3]; 8.3% scoring ≥ 80) (Dunn's test: p = 0.006 and p = 0.013, respectively).

Self-reported confidence in clinical work was also positively associated with ÇADS performance (Spearman $r_s = 0.205$, p = 0.019). Trainees with very high self-confidence (score 5) had a median ÇADS score of 3.5 [IQR: 3-4], with 50.0% achieving scores ≥ 80, whereas those with very low confidence (score 1) had a median score of 2 [IQR 2-3], with none scoring ≥ 80.

No significant associations were found between ÇADS performance and study frequency outside clinical hours (p = 0.644) or the use of specific study techniques such as spaced repetition, note-taking, or case-based exercises (all p > 0.19).

To assess whether the observed associations between AI use and ÇADS performance might be confounded by other study-related variables, we examined the correlations between the key predictors. No significant correlation was found between AI use frequency and study frequency outside clinical hours ($r_s = 0.067$, p = 0.421), between AI use and self-confidence level ($r_s = 0.124$, p = 0.138), or between

Table 3. Study habits, examination participation, and score distribution (n = 148).

Variable	n	%
ÇADS participation		
Regular attendance	98	66.2
Occasional	32	21.6
Never	18	12.2
Most recent ÇADS score		
80-100	26	20.0
60-79	95	73.1
40-59	7	5.4
<40	2	1.5
Participants who had not taken ÇADS (n = 18) were excluded from this distribution		
EDAIC participation		
Have not sat yet	142	95.9
Sat but did not pass at least one part	4	2.7
Sat and passed	2	1.4
Preferred study resources (multiple response)		
Textbooks	137	92.6
Practice question banks	12	8.1
Educational videos	63	42.6
Online articles/journals	78	52.7
AI tools	47	31.8
Group study	4	2.7
Study techniques used (multiple response)		
Question solving	13	8.8
Spaced repetition	90	60.8
Note-taking	81	54.7
Case-based exercises	48	32.4
Flashcards/post-it notes	18	12.2
Group study	7	4.7

AI: Artificial intelligence, ÇADS: *Çalışan Asistan Değerlendirme Sınavı* (trainee assessment examination conducted twice yearly by TARD), EDAIC: European Diploma in Anaesthesiology and Intensive Care. Multiple-response items sum to more than 100%.

Percentages for ÇADS score categories were calculated among participants who had completed at least one ÇADS examination (n = 130).

self-confidence and study frequency ($r_s = 0.089$, $p = 0.285$). These exploratory analyses showed that both frequent AI use and higher self-confidence were associated with higher ÇADS score categories. However, because no multivariable adjustment was performed, these findings should be interpreted as exploratory associations rather than evidence of independent effects and require confirmation in future prospective studies using adjusted statistical models.

DISCUSSION

This national survey of anesthesiology trainees in Türkiye revealed three key findings: first, that self-reported study outside clinical

hours is infrequent and brief, with the majority of respondents citing fatigue and workload as primary barriers; second, that AI tools are perceived as useful but remain underutilized, with only one in four trainees reporting regular use; and third, that frequent AI use and higher self-confidence were associated with higher ÇADS score categories in exploratory univariable analyses, whereas traditional study habits showed no such association.

The finding that 60% of respondents study only “occasionally” and that typical study sessions last one to two hours or less aligns with previous reports describing limited and irregular self-directed learning among medical trainees under high clinical workloads (15). In a survey of radiology trainees in the United States, Toy et al. (1) similarly reported that most trainees studied less than two hours per day, with time constraints cited as the primary limiting factor. Likewise, studies from Japan and other countries have documented inverse associations between clinical duty hours and available study time, suggesting that service demands directly compete with educational activities (5,16).

What distinguishes our cohort is the prominence of fatigue and burnout over time constraints alone. Nearly half of the respondents (46.9%) identified fatigue and burnout as the main barriers to increased study, a higher proportion than those citing time limitations due to clinical duties (29.4%). This pattern may reflect the intense on-call structure and unpredictable workload characteristics of anesthesiology training in Türkiye (15). Previous work has shown that burnout is associated not only with reduced study time but also with lower motivation, impaired concentration, and decreased long-term retention (16). Notably, this pattern differs from recent findings in countries with regulated duty hours. Following the implementation of work-hour restrictions in the United States, personal and family responsibilities, rather than fatigue, emerged as the primary barrier to study time among trainees (17). Nearly all respondents (96.6%) reported needing to study more, despite minimal actual study, suggesting a cycle in which workload and exhaustion reduce both the capacity and the effectiveness of self-directed learning.

Although two-thirds of the respondents in the present study rated AI tools as useful or very useful, only 24% reported using them regularly (2-3 times per week or more) for study purposes, and 38% reported never using them at all. This gap between perceived utility and actual use is consistent with early adoption patterns observed in other educational technologies.

Several factors may contribute to this discrepancy. First, moderate reliability scores (mean 3.2 out of 5) suggest that trainees remain cautious about the accuracy of AI-generated content. Concerns about factual errors, outdated information, or inappropriate clinical recommendations in LLM outputs have been widely documented (10,11,18,19). Although tools such as ChatGPT have demonstrated passing-level performance on components of medical licensing examinations, performance on structured test questions does not guarantee reliability in open-ended clinical reasoning or nuanced case interpretation (11,12). Trainees may be understandably hesitant to rely on AI tools for high-stakes learning without external validation.

Second, a lack of familiarity or institutional guidance may limit uptake. Unlike traditional textbooks or peer-reviewed journals, AI tools, such as ChatGPT, do not come with established protocols

Table 4. Factors associated with ÇADS score: Comparative analysis (n = 130).

Variable/comparison	n	Median [IQR] ÇADS score [†]	% scoring 80-100	p-value ^a
A. Frequency of AI tool use (Kruskal-Wallis)				
Daily*	12	3 [3-4]	33.3	0.0034
2-3 times/week**	19	3 [3-4]	42.1	
Once a week	8	3.5 [3-4]	50.0	
Occasionally	43	3 [3-3]	14.0	
Never	48	3 [3-3]	8.3	
B. Study frequency outside clinical hours (Kruskal-Wallis)				
Every day	5	3 [3-3]	20.0	0.6443
2-3 times/week	25	3 [3-3]	28.0	
Once a week	22	3 [3-3]	13.6	
Occasionally	78	3 [3-3]	19.2	
C. Study technique use (Mann-Whitney U per technique)				
Technique	Used n	Median (users vs. non-users)	% scoring 80-100 (users)	p-value ^b
Question solving	13	3.0 vs. 3.0	30.8	0.1900
Spaced repetition	78	3.0 vs. 3.0	17.9	0.4240
Case-based exercises	42	3.0 vs. 3.0	26.2	0.1959
Note-taking	73	3.0 vs. 3.0	20.5	0.6544
Flashcards	17	3.0 vs. 3.0	29.4	0.6498
Group study	7	3.0 vs. 3.0	42.9	0.3801
D. Self-confidence level × ÇADS score (Spearman correlation)				
Self-confidence level	n	Median ÇADS score	% scoring 80-100	p-value ^c
Very low (1)	9	2 [2-3]	0.0	0.0194
Low (2)	14	3 [3-3]	21.4	
Moderate (3)	65	3 [3-3]	13.8	
High (4)	34	3 [3-4]	29.4	
Very high (5)	8	3.5 [3-4]	50.0	

^a: ÇADS scores encoded as ordinal ranks: 1 = < 40, 2 = 40-59, 3 = 60-79, 4 = 80-100. Median [IQR] scores represent the central rank and IQR of each group. Bold p-values indicate statistical significance (p < 0.05).

^a: Kruskal-Wallis test, ^b: Mann-Whitney U test, ^c: Spearman rank correlation coefficient.

Post-hoc comparison of AI use frequency and ÇADS score (Dunn's test):

- Daily vs. never: p = 0.006*,

- 2-3/week vs. never: p = 0.013**.

No significant association was found between study technique, study frequency, institution type, or perceived barrier and ÇADS performance.

IQR: Interquartile range, AI: Artificial intelligence, ÇADS: *Çevrimiçi Asistan Değerlendirme Sınavı*.

for use in medical training. Trainees may be uncertain about how to integrate these tools effectively into their study routines or may lack awareness of “writing effective prompts,” fact-checking, and source verification. Educational programs that provide training on the responsible use of AI tools—emphasizing both their strengths and their limitations—may help bridge this gap (20).

Exploratory univariable analyses demonstrated an association between frequent AI use and higher ÇADS score categories. Trainees who used AI tools daily or several times per week were significantly more likely to achieve scores in the top performance category (≥ 80) compared to non-users (33-42% vs. 8%, p < 0.01). Although AI use frequency was not significantly correlated with study frequency or self-confidence, these exploratory analyses do not establish an independent effect and should not be interpreted causally.

Several mechanisms may explain this finding. First, AI tools such as ChatGPT allow for rapid, on-demand retrieval of explanations, definitions, and summaries, which may facilitate just-in-time learning during clinical rotations when formal study time is scarce (21,22). Unlike textbooks, which require sustained reading, or online searches, which yield variable-quality results, conversational AI can provide immediate, synthesized responses tailored to specific queries. This may be particularly valuable in anesthesiology, where trainees encounter diverse clinical scenarios requiring the rapid integration of pharmacology, physiology, and procedural knowledge.

Second, interactive questioning through AI platforms may support active learning and self-testing, both of which are associated with improved retention compared to passive reading (23).

Third, it is possible that frequent AI users represent a subset of trainees who are generally more engaged in self-directed learning, more comfortable with digital tools, or more motivated to seek out supplementary resources. However, the lack of correlation between AI use and traditional study frequency ($r_s = 0.07$, $p = 0.42$) makes a simple selection effect less likely. Instead, AI use appears to function as a distinct learning behavior independent of conventional study habits.

However, this association should be interpreted cautiously. Because of the cross-sectional design, the direction of the relationship cannot be determined. Frequent AI use may contribute to learning; however, the reverse relationship is also plausible. Trainees with higher examination performance may be more motivated, more digitally literate, or more capable of using AI tools effectively and critically. Therefore, AI use should not be interpreted as a direct cause of better examination performance in the present study.

Nonetheless, the magnitude of the observed association suggests that the relationship warrants further investigation through prospective or interventional designs.

Self-reported confidence in clinical work showed a modest but statistically significant positive correlation with ÇADS performance ($r_s = 0.21$, $p = 0.02$). This finding is consistent with previous studies linking self-efficacy to academic achievement in medical training (4,24). Trainees with higher confidence may be more likely to engage in challenging study activities and perform better under examination conditions.

However, the direction of causality remains unclear. Confidence may drive better performance, or successful performance may reinforce confidence, creating a positive feedback loop. Alternatively, both may be influenced by underlying factors such as prior academic preparation, clinical exposure, or quality of mentorship. The lack of correlation between confidence and study frequency in our data ($r_s = 0.09$, $p = 0.29$) suggests that confidence does not simply reflect greater study effort but may instead relate to clinical experience, teaching quality, or individual psychological traits (24).

Study Limitations

Several limitations should be acknowledged. First, the cross-sectional design precludes causal inference. The observed association between AI use and examination performance may reflect reverse causation, unmeasured confounding, or both. Second, all data were self-reported and therefore subject to recall bias and social desirability bias. ÇADS performance was based on self-reported score categories rather than verified official examination records. This may have introduced recall bias, selective reporting, or social desirability bias, particularly if trainees with higher scores were more willing to report their performance. Trainees may have overestimated their study time, underreported barriers, or selectively recalled AI use patterns. Third, ÇADS scores were available for only 130 of the 148 respondents, and performance was self-reported rather than verified through official records. Fourth, the survey used a convenience sample recruited through professional communication networks, which may have resulted in selection bias toward trainees who were more engaged, digitally active, or interested in AI-related topics. Because the total number of eligible trainees who received or

viewed the survey invitation was unknown, a formal response rate could not be calculated. Consequently, non-response bias cannot be excluded, and the findings may not be fully representative of all anesthesiology trainees in Türkiye. Fifth, we did not assess specific patterns of AI use—such as the types of prompts used, the depth of engagement, or whether trainees verified AI-generated responses—limiting our ability to explain why frequent use was associated with better performance. Sixth, the questionnaire was developed specifically for this study and, although it underwent internal review and pilot testing, it was not a formally validated instrument. Finally, generalizability to other specialties, countries, or training systems may be limited. Because no multivariable analysis was performed, potential confounding factors could not be adjusted for, and the observed associations should therefore be interpreted with caution.

Our findings suggest that anesthesiology training programs in Türkiye should address workload and burnout as primary obstacles to effective learning, rather than focusing solely on providing more educational content. Interventions might include protected study time, wellness initiatives, or restructured duty schedules that prioritize recovery and sustained attention (25).

Regarding AI tools, the gap between perceived usefulness and actual use indicates a need for structured guidance. Training programs could introduce formal sessions on how to use AI tools effectively and safely, emphasizing prompt design, fact-checking, and appropriate use cases.

Future research should employ prospective or interventional designs to determine whether encouraging AI use directly improves learning outcomes or whether the observed association is driven by learner characteristics. Studies should also explore which specific patterns of AI use are most beneficial, how to mitigate reliability concerns, and whether AI tools can be integrated into formal curricula without displacing evidence-based learning strategies such as spaced repetition and active retrieval practice.

CONCLUSION

In conclusion, anesthesiology trainees in Türkiye report irregular and brief study patterns, with fatigue and clinical workload cited as the primary barriers to increased self-directed learning. Although AI tools are perceived as useful, adoption remains limited, likely due to moderate trust and a lack of guidance. Frequent AI use and higher self-confidence were associated with better performance in national in-training examinations, suggesting that these factors may be associated with better examination performance under time-constrained training conditions. Educational interventions aimed at reducing burnout while providing structured guidance for the responsible use of AI tools may improve learning environments for anesthesiology trainees.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Gazi University Ethics Commission (approval no: 13.01.2025-E.1141829, date: 13.01.2025).

Informed Consent: Participation was voluntary and anonymous, and all respondents provided electronic informed consent before proceeding to the survey.

Acknowledgments

Beyza Büyükgöbüz Yeşil participated in the ESAIC Mentorship Programme as a mentee (ID: ESAIC_MSP_2024_BB).

Footnotes

Authorship Contributions

Concept: Ç.Ö., Design: Ç.Ö., Data Collection or Processing: M.G.Ö., Analysis or Interpretation: B.B.Y., Literature Search: B.B.Y., Writing: Ç.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Toy D, Escalon JG, Groner LK, Naeger DM. Preparation for the ABR core exam: resident study habits and the value of case conferences. *J Am Coll Radiol*. 2021; 18: 615-9.
2. Liu A, Reed S, Mahan JD, Wallihan R. Exploring pediatric resident attitudes and preferences for board exam preparation. *Cureus*. 2020; 12: e8022.
3. Velez DR. Modern didactic formats in surgery: a systematic review. *Am Surg*. 2023; 89: 1701-8.
4. Barragán-Hervella RG, Gaytán-Fernández S, Ramírez-Polanco A, Martínez-Asención JP, Castillo-Ramírez IO, Tapia-Cortes C, et al. Study habits and academic performance in residents of orthopedics and family medicine. *Rev Med Hosp Gen Mex*. 2023; 86: 4-11.
5. Nagasaki K, Nishizaki Y, Shinozaki T, Kobayashi H, Tokuda Y. Association between resident duty hours and self-study time among postgraduate medical residents in Japan. *JAMA Netw Open*. 2021; 4: e210782.
6. Rao P, Jordan K, Burrows H, Helms L, Latch R, Perin R, et al. Association between in-training exam scores and clinical exposure during the COVID-19 pandemic. *Acad Pediatr*. 2024; 24: 3-7.
7. Wilf-Miron R, Kagan I, Saban M. Health behaviors of medical students decline towards residency: how could we maintain and enhance these behaviors throughout their training. *Isr J Health Policy Res*. 2021; 10: 13.
8. Subramani M, Jaleel I, Krishna Mohan S. Evaluating the performance of ChatGPT in medical physiology university examination of phase I MBBS. *Adv Physiol Educ*. 2023; 47: 270-1.
9. Khan RA, Jawaid M, Khan AR, Sajjad M. ChatGPT - Reshaping medical education and clinical management. *Pak J Med Sci*. 2023; 39: 605-7.
10. Lee H. The rise of ChatGPT: Exploring its potential in medical education. *Anat Sci Educ*. 2024; 17: 926-31. Erratum in: *Anat Sci Educ*. 2024; 17: 1779.
11. Gilson A, Safranek CW, Huang T, Socrates V, Chi L, Taylor RA, et al. How does ChatGPT perform on the United States Medical Licensing Examination (USMLE)? The implications of large language models for medical education and knowledge assessment. *JMIR Med Educ*. 2023; 9: e45312. Erratum in: *JMIR Med Educ*. 2024; 10: e57594.
12. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLOS Digit Health*. 2023; 2: e0000198.
13. Li R, Kumar A, Chen JH. How chatbots and large language model artificial intelligence systems will reshape modern medicine: fountain of creativity or Pandora's box? *JAMA Intern Med*. 2023; 183: 596-7.
14. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (CHERRIES). *J Med Internet Res*. 2004; 6: e34. Erratum in: doi:10.2196/jmir.2042.
15. Yakar MN, Bıyıkoglu AS, Erbay Z, Öztürk MC, Öner Ö, Küçük M, et al. The effect of improvements in monthly income, employment conditions, and working hours on quality of life, anxiety, and burnout among anesthesiology residents: a before-after study. *J Clin Pract Res*. 2024; 46: 325-32.
16. Fafard J, Snell L. Reading habits of house-staff: what, where and why. *Med Teach*. 1989; 11: 279-83.
17. Edson RS, Beckman TJ, West CP, Aronowitz PB, Badgett RG, Feldstein DA, et al. A multi-institutional survey of internal medicine residents' learning habits. *Med Teach*. 2010; 32: 773-5.
18. Çamlar M, Sevgi UT, Erol G, Karakaş F, Doğruel Y, Güngör A. Comparative performance of neurosurgery-specific, peer-reviewed versus general AI chatbots in bilingual board examinations: evaluating accuracy, consistency, and error minimization strategies. *Acta Neurochir (Wien)*. 2025; 167: 241.
19. Waldock WJ, Zhang J, Guni A, Nabeel A, Darzi A, Ashrafian H. The accuracy and capability of artificial intelligence solutions in health care examinations and certificates: systematic review and meta-analysis. *J Med Internet Res*. 2024; 26: e56532.
20. Xu Y, Jiang Z, Ting DSW, Kow AWC, Bello F, Car J, et al. Medical education and physician training in the era of artificial intelligence. *Singapore Med J*. 2024; 65: 159-66.
21. Mondal H, Karri JKK, Ramasubramanian S, Mondal S, Juhi A, Gupta P. A qualitative survey on perception of medical students on the use of large language models for educational purposes. *Adv Physiol Educ*. 2025; 49: 27-36.
22. Banerjee M, Chiew D, Patel KT, Johns I, Chappell D, Linton N, et al. The impact of artificial intelligence on clinical education: perceptions of postgraduate trainee doctors in London (UK) and recommendations for trainers. *BMC Med Educ*. 2021; 21: 429.
23. Luo X, Alias BS, Adnan NH. Active learning strategies and student engagement in undergraduate medical education: a systematic review (2015-2025). *Anat Sci Educ*. 2026; 19: 1157-69.
24. Patenaude B, Hogan RE, Manguvo A. Exploring the dynamics of medical students' exam performance in relation to study habits, growth mindset, confidence levels, and demographics. *Med Sci Educ*. 2024; 34: 371-8.
25. Martinelli SM, Tran TN, Canter CL, Hu D, Isaak RS, Afroze F, et al. Educating anesthesiology residents and their support persons to improve understanding and increase wellness: a multisite study. *Anesth Analg*. 2026.