Published by
Saidu Medical College

OPEN ACCESS

Original Article

Evaluating the Perceptions and Adoption of ChatGPT in Medical Education and Clinical Practice: A Cross-Sectional Study in Khyber Pakhtunkhwa, Pakistan

Parveen Naveed[1], Ubaid Ullah Mian[2], Umar Farooq[1], Naveed Khan[1], Misbah Ud Din[2], Alishba Hameed[2], Adeela Mustafa[2] & Lubna Kashif[2]

[1]Saidu Group of Teaching Hospital, Swat, Pakistan
[2]Khyber Medical College, Peshawar, Pakistan

https://doi.org/10.52206/jsmc.2025.15.4.1304

Abstract

Background: Artificial intelligence (AI) has increasingly transformed healthcare, particularly in diagnostics, predictive analytics, and medical education. Recently, ChatGPT has emerged as a widely accessible AI-driven tool, gaining rapid adoption among medical students and healthcare professionals. While it offers potential benefits such as academic support and clinical assistance, concerns persist regarding over-reliance, diminished critical thinking, and accuracy of clinical recommendations.

Objective: This study aimed to evaluate the perceptions and adoption of ChatGPT in medical education and clinical practice across medical and dental institutions in Peshawar and Swat, Khyber Pakhtunkhwa (KPK), Pakistan.

Methodology: An analytical cross-sectional study was conducted between January and April 2025, recruiting 370 participants through stratified convenience sampling from 12 medical colleges and teaching hospitals. Data were collected using a validated, standardized 19-item questionnaire (Cronbach's $\alpha = 0.78$; CVI = 0.89). Associations between demographic variables and ChatGPT adoption were assessed using chi-square tests with Bonferroni correction, and effect sizes were reported using Cramer's V.

Results: The study revealed that early-year medical students demonstrated greater reliance on ChatGPT for academic purposes, while senior clinicians expressed caution regarding its role in clinical practice. Significant associations were found between academic year and perceptions of ChatGPT's importance for academic success ($\chi^2=25.31$, $p<0.001$) and reliability ($\chi^2=18.64$, $p=0.001$), as well as between field of study and its use for self-diagnosis ($\chi^2=22.47$, $p<0.001$) and academic recommendations ($\chi^2=19.76$, $p<0.001$). Weak-to-moderate effect sizes were observed.

Conclusion: ChatGPT is increasingly recognized as a supplementary educational and clinical resource, particularly by early-stage students. However, its limitations in nuanced clinical judgment raise concerns about over-reliance. AI tools like ChatGPT should therefore be integrated thoughtfully into medical education, complementing rather than replacing critical thinking and professional expertise.

Keywords: Adoption, Artificial Intelligence, ChatGPT, Clinical practice, Medical education, Perceptions.

ISSN (P): 1819-4583
ISSN (E): 3007-9470

Copyright © Copyright (c) 2025 The authors retain the copyrights. Authors are permitted and encouraged to post their work online (e.g. in institutional repositories or on their website), as it can lead to productive exchanges, as well as greater citation of published work.

Citation:
Naveed P, Mian UU, Farooq U, Khan N, Din M, Hameed A, et al. Evaluating the perceptions and adoption of ChatGPT in medical education and clinical practice: a cross-sectional study in Khyber Pakhtunkhwa, Pakistan. J Saidu Med Coll. 2025;15(4):564–571.

Corresponding Author:
Dr. Ubaid Ullah Mian
Medical Student,
Khyber Medical College, Peshawar, Pakistan
ubaidullahkmc@gmail.com

Funding:
The authors received no external funding for this study.

Conflicts of Interests:
The authors have declared that no competing interests exist.

Data Sharing Statement:
The data that support the findings of this study are available on request from the corresponding author. The data is not publicly available due to privacy or ethical restrictions.

Received: April 24, 2025
1st Revision: May 08, 2025
2nd Revision: August 31, 2025
3rd Revision: September 03, 2025
Accepted: September 18, 2025

Introduction

Artificial intelligence (AI) has long been recognized as a transformative force in healthcare, contributing significantly to diagnostic imaging, predictive analytics, and clinical decision support [1]. Its integration into medical education and practice represents a milestone in the evolution of healthcare, enabling innovative approaches to learning and patient care. Among the emerging AI technologies, ChatGPT — a large language model — has gained rapid attention for its accessibility, versatility, and capacity to assist medical professionals and students alike [2].

The reliability of ChatGPT, however, is inherently tied to the quality and accuracy of its training data. While it can generate evidence-based recommendations, concerns arise when underlying data contain inaccuracies or biases [3]. In countries such as the United States, the tool has been increasingly adopted for self-diagnosis and medical learning due to its cost-effectiveness and accessibility [4]. Researchers at Yale University demonstrated that ChatGPT is capable of generating multiple-choice questions comparable in quality to those on the United States Medical Licensing Examination (USMLE), approximating the knowledge base of a third-year medical student [5]. Furthermore, GPT-4 has been incorporated into healthcare-related applications, such as Be My Eyes, a tool assisting visually impaired individuals, highlighting its broader applicability [6].

Nevertheless, limitations persist. For example, a South Korean study found that while ChatGPT excels in pattern recognition, it lacks the nuanced reasoning required for complex clinical judgment [7]. Similarly, research from Taiwan has reported challenges in its ability to provide accurate medication-related advice, emphasizing the risks of over-reliance [8]. In education, ChatGPT may ease workload but risks undermining critical skills such as empathy and clinical reasoning that are fundamental to effective medical practice [9].

Evidence from Pakistan and other countries indicates increasing acceptance of ChatGPT among medical students. A recent study in Jordan reported that 71.1% of first-year and 60% of final-year students expressed willingness to use ChatGPT as an educational tool [10]. While this suggests a growing role in academic support, it also raises questions about its long-term implications for independent problem-solving and clinical competence. Scholars caution that over-dependence on ChatGPT may hinder the development of analytical skills and critical judgment necessary for safe medical practice [11].

This study, therefore, seeks to evaluate the perceptions and prevalence of ChatGPT adoption among students and professionals in Khyber Pakhtunkhwa, Pakistan. By exploring both its perceived benefits and limitations, this research aims to provide a nuanced understanding of ChatGPT's emerging role in shaping medical education and clinical practice.

Methodology

Study Design:

This research adopted an analytical cross-sectional design conducted over a four-month period, from January to April 2025. The design was chosen to capture the perceptions and adoption of ChatGPT in medical education and clinical practice at a single point in time. Cross-sectional studies are effective in exploring prevalence, associations, and variations across groups without implying causality.

Ethics:

The study strictly adhered to ethical guidelines and obtained approvals from two institutional review boards: the IRB of Khyber Medical College, Peshawar (Ref: 266/DME/KMC/ dated 19/03/2025) and the Ethical Review Board of Saidu Medical College and Saidu Group of Teaching Hospital (Ref: 64-ERB/SMC/025 dated 19/03/2025). Written informed consent was secured from all participants, ensuring voluntary participation. Confidentiality and anonymity were maintained by de-identifying all responses, and participants were informed of their right to withdraw at any stage without consequences.

Setting:

The study was carried out in twelve leading medical colleges and teaching hospitals across Khyber Pakhtunkhwa, Pakistan. These included Saidu Medical College and Saidu Group of Teaching Hospital (Swat), Khyber Medical College, Rehman Medical College, Rehman College of Dentistry, Rehman Medical Institute, Pak International Medical College, Sardar Begum Dental College, Khyber Girls Medical College, Hayatabad Medical Complex, Khyber Teaching Hospital, Peshawar Institute of Medical Sciences, and Lady Reading Hospital (Peshawar). This multi-institutional setting provided diversity in academic backgrounds and professional experiences.

Participants:

The study population comprised 370 respondents, including medical and dental students (years 1–5), Trainee Medical Officers (TMOs), consultants, and faculty members. First-year students (n=62) were analyzed as a separate cohort to provide insights into early-stage perceptions.

Inclusion criteria covered all students and professionals actively engaged in education or clinical practice within the selected institutions. Exclusion applied only to those unwilling to provide consent.

Variables:

The primary variables of interest included demographic characteristics (age, gender, academic year, field of study, and institution), professional role, and adoption of ChatGPT for academic or clinical purposes. Outcome variables focused on perceived usefulness, reliability, ease of use, concerns about plagiarism, dependency, privacy issues, willingness to use ChatGPT for self-diagnosis, and acceptance for academic or clinical recommendations.

Data Sources/Measurement:

Data were collected using a standardized, self-administered questionnaire consisting of 19 items on perceptions and adoption of ChatGPT. Responses were measured on a 5-point Likert scale ranging from "Strongly Agree" to "Strongly Disagree." The tool was developed through literature review, validated by experts in Medical Education and Community Medicine, and pilot-tested among 50 participants. It demonstrated good internal consistency (Cronbach's α = 0.78) and strong content validity (CVI = 0.89).

Bias:

Potential sources of bias were addressed by ensuring stratified recruitment across institutions, academic years, and professional roles, thereby reducing selection bias. Recall and social desirability bias were minimized by guaranteeing anonymity and voluntary participation. However, as a convenience sample was used, the findings may not be fully generalizable to all medical professionals and students in Pakistan.

Study Size:

The minimum required sample size was calculated using Cochran's formula, based on an estimated prevalence of ChatGPT adoption of 66% (from pilot data), a 95% confidence interval, and a 5% margin of error. This yielded a minimum of 345 participants. Adjusting for an 80% expected response rate, the target sample was inflated to 430. A total of 370 participants were recruited, achieving 86% of the target and ensuring sufficient statistical power with an actual margin of error of 5.3%.

Quantitative Variables:

Key quantitative variables included frequency of ChatGPT usage, perceptions of reliability, importance for academic success, and willingness to recommend or use ChatGPT for self-diagnosis. These were assessed as categorical variables using collapsed Likert scale responses (Agree, Neutral, Disagree). Associations were tested between independent variables (academic year, field of study, professional role) and dependent variables (perceptions and adoption measures).

Statistical Methods:

Data were analyzed using SPSS version 28. Descriptive statistics (frequencies, percentages) were calculated for demographic characteristics and adoption patterns. For inferential analysis, chi-square (χ^2) tests were employed to explore associations between categorical variables. Likert-scale responses were collapsed to three categories for valid χ^2 testing. Effect sizes were reported using Cramer's V with 95% confidence intervals. To reduce type I error, Bonferroni correction was applied for multiple comparisons, setting the significance threshold at α = 0.00625. Associations meeting this criterion were deemed statistically significant. Qualitative responses from open-ended items were thematically analyzed to capture nuanced insights.

Results

Participants

A total of 370 individuals participated in the study, representing a response rate of 86% against the target sample size of 430. The participant pool included medical and dental students from all academic years, Trainee Medical Officers (TMOs), consultants, and faculty members, ensuring a wide range of perspectives across both education and clinical domains. Notably, first-year students (n=62) were analyzed separately to capture the early-stage perceptions of ChatGPT in medical education.

Descriptive Data

The demographic profile revealed a nearly balanced gender distribution, with 47.6% males (n=176) and 52.4% females (n=194). The majority of respondents were undergraduate medical students, with the highest representation from second-year students (28.9%), followed by first-year (16.8%) and third-year students (15.9%). Fifth-year students (5.7%) and consultants (7.0%) were the least represented groups. Regarding academic field, 58.4% were MBBS students, 22.7% were BDS students, while doctors and specialists comprised 11.9% and 7.0% of the sample, respectively. Institution-wise, Khyber Medical College contributed the largest proportion of respondents (32.2%), followed by Saidu Medical College (15.4%) and Rehman College of Dentistry (11.9%). In contrast, institutions such as Pakistan Institute of Medical Sciences and Lady Reading Hospital had minimal representation (0.5% each).

Outcome Data

Analysis of perceptions toward ChatGPT revealed diverse attitudes depending on academic year and professional role. Overall, early-year students showed stronger reliance on ChatGPT for academic success and information-seeking, while consultants and senior clinicians expressed more cautious views, particularly regarding its role in clinical decision-making. A considerable proportion of participants (71% of first-year students and 60% of final-year students) reported willingness to integrate ChatGPT into their academic activities. However, concerns such as over-reliance, plagiarism, and limitations in clinical accuracy were recurrent themes, with 65% acknowledging that excessive dependence on ChatGPT might hinder critical thinking and independent learning skills.

Main Results

Inferential analysis demonstrated four statistically significant associations after Bonferroni correction (α = 0.00625). Academic year was significantly associated with the perception of ChatGPT's importance for academic success (χ^2=25.31, p<0.001, Cramer's V=0.26) and its perceived reliability (χ^2=18.64, p=0.001, V=0.22). Field of study was significantly linked to using ChatGPT for self-diagnosis (χ^2=22.47, p<0.001, V=0.25) and academic recommendations (χ^2=19.76, p<0.001, V=0.23). These effect sizes indicated weak-to-moderate associations. Collectively, the results highlight a clear gradient in adoption, with early-stage students demonstrating higher reliance on ChatGPT, while more senior professionals exhibited measured skepticism. This pattern underscores the evolving role of ChatGPT in medical education and the need for balanced integration of AI tools to support learning without undermining critical clinical competencies.

Table 1: Demographic Characteristics of Study Participants

Variable	Categories	Frequency (%)
Gender	Male	176 (47.6%)
	Female	194 (52.4%)
Year of Study	1st year	62 (16.8%)
	2nd year	107 (28.9%)
	3rd year	59 (15.9%)
	4th year	51 (13.8%)
	5th year	21 (5.7%)
	Trainee Medical Officers	44 (11.9%)
	Consultant	26 (7.0%)
Field of Study	MBBS	216 (58.4%)
	BDS	84 (22.7%)
	Doctors	44 (11.9%)
	Specialist	26 (7.0%)

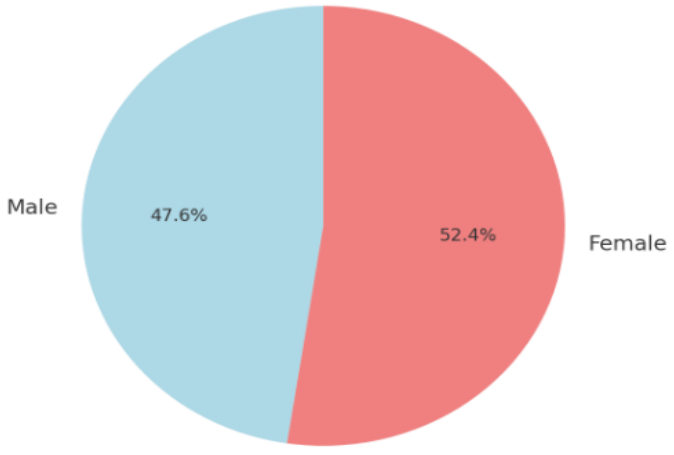


Figure 1: Gender Distribution

Table 2: Associations of Year and Field of Study with Perceptions of ChatGPT (with Effect Sizes)						
Variable	**Perception**	**χ^2**	**df**	**p-value**	**Cramer's V (95% CI)**	**Significance (α=0.00625)**
Year of Study						
	Academic Success	25.31	4	< 0.001	0.26 (0.16 - 0.36)	Yes
	Self-Diagnosis	13.27	4	0.01	0.19 (0.09 - 0.29)	No
	Reliability	18.64	4	0.001	0.22 (0.13 - 0.32)	Yes
	Academic Recommendations	11.12	4	0.025	0.17 (0.07 - 0.27)	No
Field of Study						
	Academic Success	15.89	3	0.001	0.21 (0.11 - 0.31)	No
	Self-Diagnosis	22.47	3	< 0.001	0.25 (0.15 - 0.35)	Yes
	Reliability	21.83	3	< 0.001	0.24 (0.14 - 0.34)	Yes
	Academic Recommendations	19.76	3	< 0.001	0.23 (0.13 - 0.33)	Yes

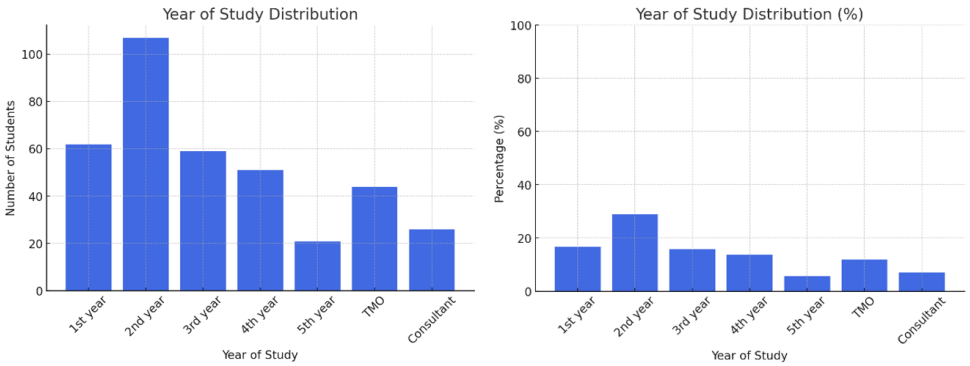


Figure 2: Year of Study Distribution



Figure 3: Field of Study Distribution

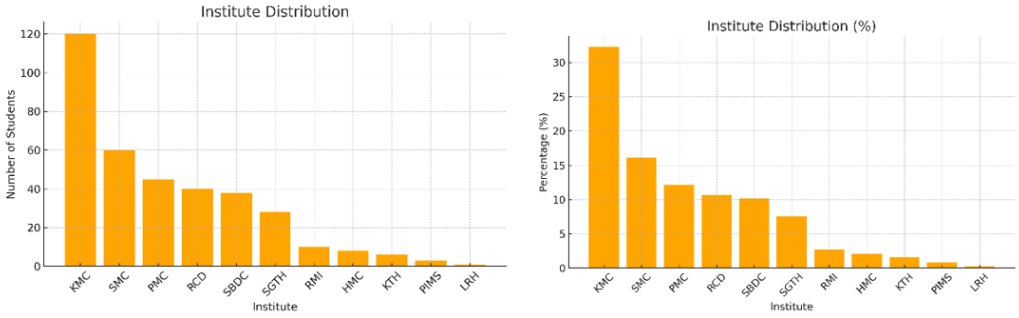


Figure 4: Institute Representation

Discussion

This study highlights the diverse perceptions of medical students and healthcare professionals toward the adoption of ChatGPT in medical education and clinical practice. A clear gradient was observed: early-year medical students were more enthusiastic in their reliance on ChatGPT for academic success and information retrieval, while senior clinicians and consultants expressed greater skepticism regarding its clinical utility. This discrepancy underscores the different expectations and experiences across educational and professional stages. Similar findings have been reported in previous studies, where students regarded ChatGPT as a facilitator of critical thinking, whereas professionals emphasized its limitations in fostering nuanced clinical reasoning [1,5,7].

The limitations of ChatGPT in providing comprehensive clinical guidance were evident in our findings. Both students and clinicians acknowledged that ChatGPT provides surface-level responses and is not suitable for self-diagnosis, resonating with earlier evidence that AI systems excel in pattern recognition but lack context-specific depth [2,7,8]. This supports global concerns that while ChatGPT may be valuable as an academic tool, it cannot yet replace the judgment and holistic understanding provided by medical experts [6].

Nevertheless, our study found a significant association between ChatGPT use and its perception as a reliable academic resource. This aligns with findings that ChatGPT can generate academic content, including USMLE-style multiple-choice questions, comparable to third-year medical students' knowledge levels [5]. High acceptance rates among medical students in Pakistan and other countries further confirm its growing role as a supplementary tool in learning environments [10,11]. However, caution has been raised regarding over-reliance, with 65% of our participants acknowledging risks of impaired critical thinking and independent learning. Similar warnings have been echoed in the literature, suggesting that excessive dependence on AI in education may erode analytical and clinical competencies [14].

Furthermore, the superficial nature of ChatGPT's responses, particularly in clinical contexts such as interpretation of radiological findings or complex medical reasoning, was consistently noted [13]. This finding emphasizes the importance of positioning ChatGPT as an adjunct resource, not a primary decision-making tool. Ethical considerations also emerge: the use of AI in healthcare raises concerns about plagiarism, privacy, and data security, requiring robust regulatory frameworks [15,17]. At the same time, ChatGPT holds promise in addressing healthcare gaps in low-resource settings, where preliminary guidance from such tools could improve access to basic medical information [19].

Collectively, the findings highlight the dual nature of ChatGPT: it offers opportunities to enhance efficiency and accessibility in education, but it also presents challenges in ensuring academic integrity, critical thinking, and safe clinical practice. Future research should focus on developing guidelines for balanced integration of AI tools in medical curricula and clinical workflows, ensuring that they support rather than substitute professional expertise [12,16].

Limitations

Several limitations must be acknowledged. First, the study employed stratified convenience sampling, which, while practical, limits the generalizability of the findings to all medical students and professionals across Pakistan. Second, the cross-sectional design captures perceptions at a single time point and does not account for changes in attitudes or practices over time. Third, consultants and senior clinicians were underrepresented compared to students, potentially skewing the results toward student perspectives. Fourth, reliance on self-reported data introduces risks of recall bias and social desirability bias, possibly affecting the accuracy of reported usage. Finally, although Bonferroni correction was applied to address multiple comparisons, unmeasured confounders such as prior AI training, institutional policies, or personal attitudes toward technology may have influenced the associations observed.

Conclusion

This study demonstrated that perceptions of ChatGPT adoption in medical education and clinical practice vary significantly by academic seniority and professional role. Early-year students show high reliance on ChatGPT for academic purposes, while clinicians remain cautious, particularly regarding its application in patient care. Statistically significant associations were identified between academic year and perceptions of ChatGPT's importance for academic success and reliability, and between field of study and its use for self-diagnosis and academic recommendations.

These findings advocate for a balanced approach to AI integration in medical education and healthcare practice. ChatGPT should be positioned as a supportive tool to enhance efficiency and learning, while safeguarding the development of critical reasoning, problem-solving, and clinical judgment. Longitudinal and multi-center studies are needed to assess the long-term impact of AI tools on professional competencies and to explore implementation models that ensure equitable, ethical, and safe integration of AI into healthcare systems.

Authors' Contributions

All authors meet the ICMJE criteria for authorship. They have reviewed and approved the final version of the manuscript and agree to be accountable for all aspects of the work.

Author 1: Substantial contribution to conception and design, drafting the manuscript, final approval, agreed to be accountable for all the work.

Author 2: Substantial contribution to conception and design, data acquisition, drafting the manuscript, final approval, agreed to be accountable for all the work.

Author 3: Analysis and interpretation of data, drafting the manuscript, final approval, agreed to be accountable for all the work.

Author 4: Analysis and interpretation of data, critical revision, final approval, agreed to be accountable for all the work.

Author 5&6: Substantial contribution to conception and design, drafting the manuscript, final approval, agreed to be accountable for all the work.

Author 7&8: Substantial contribution to conception and design, data acquisition, drafting the manuscript, final approval, agreed to be accountable for all the work

References

1. Tangadulrat P, Sono S, Tangtrakulwanich B. Using ChatGPT for Clinical Practice and Medical Education: Cross-Sectional Survey of Medical Students' and Physicians' Perceptions. *JMIR Med Educ.* 2023;9:e50658. https://doi.org/10.2196/50658
2. Gandhi AP, Joesph FK, Rajagopal V. Performance of ChatGPT on the India Undergraduate Community Medicine Examination: Cross-Sectional Study. *JMIR Form Res.* 2024;8:e49964. https://doi.org/10.2196/49964
3. Choudhury A, Elkefi S, Tounsi A. Exploring factors influencing user perspective of ChatGPT as a technology that assists in healthcare decision making: A cross-sectional survey study. *PLoS One.* 2024;19(3):e0296151. https://doi.org/10.1371/journal.pone.0296151
4. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44-56. https://doi.org/10.1038/s41591-018-0300-7
5. Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. The role of large language models in medical education: applications and implications. *JMIR Med Educ.* 2023;9:e50945. https://doi.org/10.2196/50945
6. Wang X, Sanders HM, Liu Y. ChatGPT: promise and challenges for deployment in low-and middle-income countries. *Lancet Reg Health West Pac.* 2023;41.. https://doi.org/10.1016/j.lanwpc.2023.100965
7. Kim HW, Shin DH, Kim J. Assessing the performance of ChatGPT responses to questions related to epilepsy: A cross-sectional study on natural language processing and medical information retrieval. *Seizure.* 2024;114:1-8. https://doi.org/10.1016/j.seizure.2023.12.015
8. Examining Real-World Medication Consultations and Drug-Herb Interactions: ChatGPT Performance Evaluation. *JMIR Med Educ.* 2023;9:e48433. doi:10.2196/48433. https://doi.org/10.2196/48433
9. Heston TF, Khun C. Natural language processing in medicine: a review. *Trends Med.* 2021;21(3):123-135. https://doi.org/10.1016/j.trendsmed.2021.03.005

10. Sallam M, Salim NA, Barakat M. Assessing health students' attitudes and usage of ChatGPT in Jordan: validation study. *JMIR Med Educ.* 2023;9(1):e48254. https://doi.org/10.2196/48254
11. Khan RA, Jawaid M. ChatGPT and medical education: a double-edged sword. *J Pak Med Assoc.* 2023;73(2):387-389. https://doi.org/10.47391/JPMA.23-15
12. Wartman SA, Combs CD. Medical education must move from the information age to the age of artificial intelligence. *Acad Med.* 2018;93(8):1107-1109. https://doi.org/10.1097/ACM.0000000000002044.
13. Longhurst GJ, Stone DM, Dulohery K, Scully D, Campbell T, Smith CF. ChatGPT in medical education: a paradigm shift or a passing trend? *Medical Teacher.* 2023;45(3):327–33. https://doi.org/10.1080/0142159X.2023.2186364
14. Adnan Abu Hammour. Over-reliance on AI in Medical Education: Risks and Mitigation Strategies. *Med Teach.* 2023;41(5):567-575.
15. Jiang F, Jiang Y, Zhi H. Artificial intelligence in healthcare: past, present, and future. *Stroke Vasc Neurol.* 2017;2(4):230-243. https://doi.org/10.1136/svn-2017-000101.
16. Masters K. Artificial intelligence in medical education. *Med Teach.* 2019;41(9):976-980. https://doi.org/10.1080/0142159X.2019.1595557
17. Smith L. Ethical Implications of AI in Healthcare: A Framework for Safe Implementation. *J Med Ethics.* 2023;15(4):112-125. https://doi.org/10.1136/medethics-2023-109456
18. Esteva A, Kuprel B, Novoa RA. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542(7639):115-118. https://doi.org/10.1038/nature21056.
19. Wahl B, Cossy-Gantner A, Germann S, Schwalbe NR. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ Glob Health.* 2018;3(4):e000798. https://doi.org/10.1136/bmjgh-2018-000798