Original Article

Building the Blueprint for Medical Educators of West Bengal: A Delphi Study on developing consensus-based competency framework for Large Language Models

Ray Soumalya[1], Roy Shuvajit[2], Bhattacharyya Agnihotri[1], Karmakar Prasanta Ray[3]

[1]Associate Professor, Department of Community Medicine, Sarat Chandra Chattopadhyay Government Medical College and Hospital, Uluberia, Howrah, West Bengal, India. [2]Senior Resident, Department of Community Medicine, Sarat Chandra Chattopadhyay Government Medical College and Hospital, Uluberia, Howrah, West Bengal, India. [3]Professor, Department of Community Medicine, Sarat Chandra Chattopadhyay Government Medical College and Hospital, Uluberia, Howrah, West Bengal, India

Abstract

Background: The increasing use of Large Language Models (LLMs) demands an official change in medical education preparation. The Indian context lacks a consensus model to inform the training of medical educators in LLM competencies. This deficit risks unprepared faculty and graduates within an AI-augmented healthcare system. The objective of this study was to develop a consensus-based list of core competencies for medical educators. **Material and Methods:** A descriptive, observational study using a modified Delphi technique was conducted over six months, April 2025 to October 2025. A twenty-five-member expert panel, purposively selected and stratified by medical education discipline, was used. All participants had considerable experience and prior AI training. Three rounds of online questionnaires were used, with the consensus for an item's inclusion was decided by a median score of 4.0 or more and an Interquartile Range (IQR) of 1.0 or less. **Results:** Core competency items showed exceptionally high levels of consensus (31/31). Thirty-one core competency items met prominent levels of consensus (IQR = 1.0 or 0.0). The competencies of greatest priority revolved around three pillars: Ethical and Safety Imperatives, Practical utility in academic roles, and a supportive Institutional Framework. **Conclusion:** This research forms a solid, consensus-built model of LLM literacy. The results provide further clear guidance for curriculum developers to shift medical education away from technical theory and toward practical skills. Ongoing curricular change and formal faculty development are required.

Keywords: Large Language Models, Medical Education, Curriculum Development, Delphi Technique, Generative AI.

Received: 23 March 2026 **Revised: 30 March 2026** **Accepted: 06 April 2026** **Published: 27 April 2026**

Introduction

Large language models (LLMs) represent a breakthrough in natural language processing by computers.[1-7] They are progressively revealing potential in medical education through syntactic reproduction and greater student understanding of medical diagnosis and physical training.[1,2-7]

Medical adoption of LLMs offers faculty and students opportunities, such as identifying curriculum deficiencies and generating learning outcomes.[1,3-10] Educators can design realistic patient scenarios to teach clinical reasoning and support the development of assessment tests and rubrics. This assistance supports the objectives of Competency-Based Medical Education.[11] LLMs serve as personalised tutors for learners and help in developing personalised study plans and immediate clarifications of complex material.[1-11] Nevertheless, the risk of algorithmic bias and ethical concerns are associated with this type of support.[3,4,8-10] Academic dishonesty presents further complications, including plagiarism and improper use in assessments.[3,4,8–10] A phenomenon known as "artificial hallucination" is another serious threat.[3,4,8] This misinformation can pose a significant risk to the patient's health in the medical field. In the history of medical education, technological changes have made social disparities more ubiquitous.[3,4,9,10] This generates a quandary in medical practice, in which doctors must ensure equitable care. The rapid integration of LLMs without rigorous oversight risks deepening existing inequities. These disparities can affect access to technology and patient outcomes.

A noticeable gap exists in the academic literature. Most studies focus on student use of LLMs or propose broad frameworks for digital health.[3,4,9,10,12] These approaches often emphasise technical skills over the specific needs of medical educators.[13] Educators interact with LLMs as sophisticated end- users, not as developers.[3,9] Their required competencies are unique. They must learn to teach with these models, use them for assessment, and critically evaluate their outputs—these needs differ from those of students.[3,4,9,12-14]

No validated curriculum currently exists to guide faculty in acquiring this skill set, particularly in the Indian context.[2] A competency-based framework is also difficult to find. The

Address for correspondence: Dr. Soumalya Ray, Associate Professor, Department of Community Medicine, Sarat Chandra Chattopadhyay Government Medical College and Hospital, Uluberia, Howrah, West Bengal, India
E-mail: drsoumalya@gmail.com

DOI:
10.21276/amit.2026.v13.i1.612

How to cite this article: Soumalya R, Shuvajit R, Agnihotri B, Ray KP. Building the Blueprint for Medical Educators of WestBengal: A Delphi Study on developing consensus-based competency framework for Large Language Models. Acta Med Int. 2026;13(1):1133-1138.

present study was designed to address this gap. The objective of the study was to develop a consensus-based list of core competencies for medical educators using the modified Delphi method.

Materials and Methods

Study Type, Design, and Duration: This was a descriptive observational study. It used a modified Delphi technique. The Delphi method was chosen to achieve expert consensus through iterative rounds of anonymous feedback. The study was conducted from April 2025 to October 2025.

Study Population and Expert Panel Selection: The study population included faculty from government medical colleges in West Bengal, India. A panel of twenty-five medical educators was approached. A stratified purposive sampling technique was used to create the panel of experts, with inclusion criteria being: being a member of a medical education unit or curriculum committee of the institute, having attended a dedicated Artificial Intelligence (AI) workshop, or having published on AI in an indexed journal in the preceding year, and who provided informed consent. The sampling strategy followed two strata: one based on geographical diversity (faculty members from the medical colleges around the state capital and the northern and western regions of the state), and the other on academic designation (faculty members from RMO/Demonstrator, Assistant Professor, Associate Professor, and Professor). It was expected to provide rich data for further analysis.

Data Collection Procedures: Data were collected through three iterative rounds of an online questionnaire. Round 1 was qualitative and focused on identifying core concepts. It used open-ended questions to gather perceptions on essential skills, topics, and ethical considerations. This round also collected demographic characteristics (it enquired about respondents' perceptions of the essential skills, topics, and ethical considerations required for educators to use LLMs effectively in teaching and research). The research team reviewed and synthesised these responses into clear competency items. Round 2 used a quantitative questionnaire with 37 distinct items. Panelists rated the importance of each item on a 5-point Likert scale for its inclusion in the list of competencies (1="Not at all Important" to 5= "Essential"). An option was available to skip items outside a panelist's expertise. Round 3 presented a refined list of 34 items that showed initial convergence. For each item, experts saw their previous rating alongside the group's median score. They were then asked to rate the items. During this round, they were also requested to share the reason for re-rating.

Data analysis: All responses were imported from Google Form to Google Sheets. After round one, responses to open-ended questions were analysed manually to identify similar words and/or phrases and combined into condensed items. The first and second authors did the analysis separately. In the event of a difference of opinion, the third author made the final decision on the suitability of the items. Demographic characteristics were analysed in Python (version 3.10). For rounds 2 and 3, quantitative analysis was performed on the data. All statistical analyses were conducted using Python (version 3.10). For each competency item, the median was used as the primary measure of central tendency to determine overall importance, and the interquartile range (IQR) was used as the measure of dispersion to quantify agreement. An a priori criterion for consensus was established: an item was included in the final framework if it achieved a median rating of 4.0 ("Very Important") or higher and an IQR of 1.0 or less.[2,15–17]

Ethical considerations: The study adhered to the ethical guidelines set forth by the Indian Council of Medical Research (ICMR).[18] Formal ethical approval was obtained from the Institutional Ethics Committee of Sarat Chandra Chattopadhyay Government Medical College and Hospital, Uluberia, Howrah. All respondents participated voluntarily; informed consent was obtained in writing (via an online form); all data privacy and security concerns were taken seriously in line with organisational and regulatory requirements; and faculty autonomy was not compromised.

Results

Twenty-five individuals were approached for participation. 19 participants responded in Round 1. 15 participants participated in both Round 2 and Round 3. The qualitative analysis of Round 1 produced 37 different compartmentalisations. After the repeated rating process, 34 items met the consensus criteria.

Characteristics of the Expert Panel: The panel of experts included professionals from various medical fields - Community Medicine (60%), Microbiology (16%), Anatomy (8%), Physiology (8%), and Forensic Medicine and Toxicology (8%). The average professional experience was 14.9 years (SD=7.1), and the median experience was 13 years. Forty percent were Professors and 40 percent were Associate Professors. The median rating for self-declared LLM proficiency was 3.0 on a 5-point scale. Most (60%) had formal training in AI, Data Science, or Health Informatics in the preceding year. [Table 1]

Overview of Delphi Rounds and Consensus: Of the 37 items rated in Round 2, 34 met the consensus criteria. 13 had an IQR of 0.0, indicating maximum agreement. This finding highlights the level of belief that the panel had in these competencies. Three items were not included in the final framework. These were: concept of Black Box, Personalised Medicine, and Longitudinal Curriculum. [Table 2 and Table 4]

Priority Items and Thematic Grouping: A list of the 13 highest-priority items showed three fundamental pillars. The first pillar emphasises Ethical and Safety Imperatives. This was the largest thematic group. This one contains competencies such as “Critical Appraisal of LLM Outputs,” “Verification of Information,” “Recognising Inappropriate Use-Cases,” “understanding Algorithmic Bias,” and ensuring “Patient Data Privacy and Security”. The second pillar is Practical Utility in Core Academic Roles. This theme covers immediate application to improve educators' day-to-day practice. At the same time, the third pillar is the requirement of a supportive Institutional Framework. This was reflected in the fact that the highest-priority items included Faculty Development Programs and Clear Institutional Guidelines. [Table 3]

Table 1: Background characteristics of respondents(n=25)

Characteristic	Frequency	Percentage
Departments		
Community Medicine	15	60
Anatomy	2	8
Physiology	2	8
Forensic Medicine and Toxicology	2	8
Microbiology	4	16
Professional Experience in Medical Education (Years)		
Mean (SD)	14.9(±7.1)	
Median (IQR)	13(7.0)	
Academic Rank		
Professor	10	40
Associate Professor	10	40
Assistant Professor	4	16
Tutor/Demonstrator	1	4
AI, Data Science, or Health Informatics Training		
Yes	15	60
No	10	40

Table 2: List of Consensus Items (n= 34)

Rank	Domain	Curriculum Item Statement	Final Median	Final IQR
1	Foundational concepts	Core Definition of AI and LLMs	5	0
1	Applications in medicine	Clinical Documentation	5	0
1	Applications in medicine	Information Synthesis	5	0
1	Applications in medicine	Medical Education and Assessment	5	0
1	Applications in medicine	Patient Communication	5	0
1	Ethical & Legal	Patient Data Privacy and Security (e.g., DPDP Act)	5	0
1	Ethical &Legal	Algorithmic bias	5	0
1	Essential Competencies	Critical appraisal of LLM Outputs	5	0
1	Essential Competencies	Verification of Information	5	0
1	Essential Competencies	Ethical Discernment	5	0
1	Essential Competencies	Recognizing Inappropriate Use-Cases	5	0
1	Pedagogy & Barriers	Faculty Development Programs	5	0
1	Pedagogy & Barriers	Clear Institutional Guidelines	5	0
2	Foundational concepts	Inherent limitations of LLMs (e.g., "hallucinations")	5	1
2	Applications in medicine	Diagnostic support	5	1
2	Applications in medicine	Research assistance	5	1
2	Ethical & Legal	Accountability and liability	5	1
2	Ethical & Legal	Informed consent for AI Systems	5	1
2	Ethical & Legal	Misinformation and disinformation	5	1
2	Ethical & Legal	Impact on the Doctor-Patient Relationship	5	1
2	Ethical & Legal	Intellectual Property and Plagiarism	5	1
2	Essential Competencies	Effective Prompt Engineering	5	1
2	Essential Competencies	Integration into workflow	5	1
2	Essential Competencies	Communicating with Patients about AI	5	1
2	Pedagogy & Barriers	Interactive Workshops	5	1
2	Pedagogy & Barriers	Case-Based Learning (CBL)	5	1
2	Pedagogy & Barriers	Online Modules	5	1
So3	Foundational Concepts	Mechanism of LLMs	4	1
3	Foundational Concepts	Data Sources and Training	4	1
3	Foundational Concepts	Evolution of LLMs	4	1
3	Ethical &Legal	Health Equity	4	1
3	Pedagogy & Barriers	Integration into Existing Courses	4	1
3	Pedagogy & Barriers	Access to technology	4	1
3	Pedagogy & Barriers	Addressing Time Constraints	4	1

Table 3: Thematic classification of Consensus Items

Core Theme	Consensus Items within Theme	Implication
Ethical & Safety Imperatives	• Patient Data Privacy and Security (e.g., DPDP Act)• Algorithmic Bias• Critical Appraisal of LLM Outputs• Verification of Information• Ethical Discernment• Recognizing Inappropriate Use-Cases	This is the most dominant theme, indicating the panel's conviction that the curriculum must be centered on the principles of responsible, safe, and critical use to mitigate risks.
Practical Utility in Core Academic Roles	• Clinical Documentation• Information Synthesis• Medical Education and Assessment• Patient Communication	This theme highlights the panel's focus on immediate, tangible applications that can enhance the efficiency and quality of an educator's daily academic and clinical tasks.
Institutional Framework	• Faculty Development Programs• Clear Institutional Guidelines	This theme underscores the panel's belief that individual competency is insufficient without structured, top-down

		support and clear policies from the institution.
Foundational Understanding	• Core Definition of AI and LLMs	This single-item theme represents the essential starting point of the curriculum, ensuring all users share a mutual understanding of the technology itself before applying it.

Table 4: List of Excluded Items (n= 3)

Item ID	Original Domain	Excluded Item Statement	Potential Rationale for Exclusion
C3	Foundational Concepts	Concept of "Black Box": Discussing the challenges in understanding the internal decision-making processes of some AI models.	This item may have been considered too abstract or theoretical for a foundational curriculum focused on practical application, with the panel prioritizing the limitations of outputs over the deep mechanics of the models.
C7	Applications in Medicine	Personalized medicine: Exploring the potential role of LLMs in tailoring patient treatment plans.	The panel likely viewed this as a future-state or highly specialized application. It may have been deprioritized in favor of more immediate, universally applicable skills relevant to all educators today.
F5	Pedagogical Approaches	Longitudinal Curriculum: A curriculum that is integrated and revisited throughout the multiple years of medical education.	While a valid concept, this likely represents a specific implementation strategy rather than a core content item. The panel may have focused on first defining what to teach before deciding how it should be integrated over time.

Discussion

The list of competencies that has resulted from this modified Delphi study is generally considered to outline the role of proactive curricular modifications to become prepared to integrate Large Language Models (LLMs) into medical education and training, albeit indirectly; however, with many healthcare professionals subtly experiencing the risks and benefits of such changes.[1,2,7,9,10,13,14,20]

Within the teaching domain, the expert consensus highlighted on the competence of customised content generation and AI-assisted interactive teaching, which directly correlates with the urgency to redesign longstanding assessment modalities that are currently susceptible to being overtaken by AI-based teaching.[1,3,4,9,14,20,21] Also, in the teaching area, it was noted that AI may overthrow the assessment and feedback systems because of its ability to provide instant and formative feedback regarding the performance of the students.

In the realm of medical research, the competencies identified, such as conducting literature reviews, writing manuscripts and editing them, data analysis, and reference management, prove the acceptance of the panel regarding the power of LLMs to make research faster through the efficient acquisition, retrieval, and synthesis of pertinent information within scientific literature and electronic medical records. To a greater extent, generative AI and LLMs are perceived as useful in addressing a range of clinical and documentation-related matters, which must include the automation of the process of documenting it, helping with clinical coding, and summarising clinical documents, etc.[20-23] The unified decision made by the panel members on these application types points to the fact that one of the areas medical educators pay their attention to is the practical value of using LLMs to become more productive and learn better in both academic and professional settings.

More importantly, the identified competencies go beyond simple use and focus on cognitive and ethical equality, suggesting apparent gaps in existing AI literacy training. The consensus placed high value on the ability to perform critical appraisal of LLM generated text and to recognise the AI-generated text.[24,25] The focus on discernment will be crucial since LLMs are prone to generate so-called 'AI hallucinations' a response that looks logical yet fails to hold. This is an issue of critical concern as it nullifies the primary benefit of learning-on-need. Physicians need to develop sufficiently important technical expertise to effectively utilise the opportunities of one of the fastest growing technical tools of medical interpretation.

Ethics and moral issues, such as legal and societal concerns, and the underlying need to act ethically when using AI technologies, emphasise ethical considerations within the panel. The absence of human interaction and emotion in LLMs is particularly critical in areas such as medical education, where empathy formation is essential.[27,29] The firm belief in ethics, especially the necessity of ethical and legal compliance when discussing health data, accountability, and traceability, aligns with the universal call for a unified ethical framework to address LLMs in medical education.[26-30]

The Delphi method used in this paper, a consensus method preferred when empirical evidence is scarce or conflicting, would work well in the new realm of the role of LLM competencies. The study has been high in consensus (34 of 37 items met the predefined consensus criteria, with extremely elevated levels of agreement, as indicated by IQRs of 1.0 or 0.0). Such a prominent level of agreement is comparable to earlier e-Delphi studies that sought to assess AI competence among medical graduates and found that participants strongly agreed on a significant list of items. The approach capitalises on the accumulated wisdom of geographically dispersed experts without undue prominence for any one person, thereby enabling strong contributions. The approach fulfilled the fundamental objective of defining competencies.[2,17,20,31,32]

The study has some limitations. The Delphi technique specifies areas of importance based on the expert panel, and it is not an objective statement of absolute truth. Another constraint is the specific geographical composition of the expert panel, which was recruited from a single region in India. While the panelists were professionally diverse, the results may not be generalizable to all medical education systems globally. Lastly, the area of GenAI is characterised by extremely rapid development. The consensus reached reflects the level of technology in existence at the time of the investigation. All competency frameworks in this field should be constantly reviewed and revised.

Conclusion

This study has developed an adequate, consensus-driven list of cores LLM pedagogies in medical educators. It offers a roadmap for incorporating the idea of generative AI (GenAI) literacy into medical education. The overall consensus on skills such as critical appraisal, ethical use, and prompt engineering indicates that they are an urgent necessity. The medical community must prioritize the practical competencies essential for future physicians to function safely in an AI-augmented healthcare environment. Judging by the list of competencies, future research on the subject should focus on specific areas. To begin with, specific educational content and syllabi implemented in medical schools must be developed in line with the identified competencies. Secondly, the effectiveness of the newly designed curriculum must be rigorously assessed. Finally, different stakeholder groups, such as students, practitioners, and regulatory experts, must be involved to test the cross-cultural applicability of these competences.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Hc L, Js U, Jr R. A systematic review of large language models and their implications in medical education.MedEduc[Internet].2024Nov[cited2025Oct24];58(11).Availablefrom: https:/pubmed.ncbi.nlm.nih.gov/38639098/
2. Developing a Canadian artificial intelligence medical curriculum using a Delphi study | npj Digital Medicine[Internet].[cited2025Oct24].Availablefrom:https:/www.nature.com/articles/s41746-024- 01307-1
3. ZhuiL,YhapN,LipingL,ZhengjieW,ZhonghaoX,XiaoshuY,et al.ImpactofLargeLanguageModels onMedicalEducationandTeachingAdaptations.JMIRMedInform.2024July25;12:e55933.
4. Abd-Alrazaq A, AlSaad R, Alhuwail D, Ahmed A, Healy PM, Latifi S, et al. Large Language Models in Medical Education: Opportunities, Challenges, and Future Directions. JMIR Med Educ. 2023 June 1;9:e48291.
5. AydinS,KarabacakM,VlachosV,MargetisK.Largelanguagemodelsinpatienteducation:ascoping reviewofapplicationsinmedicine.FrontMed.2024;11:1477898.
6. Clusmann J, Kolbinger FR, Muti HS, Carrero ZI, Eckardt JN, Laleh NG, et al. We future landscape of largelanguagemodelsinmedicine.CommunMed.2023Oct10;3(1):141.
7. Safranek CW, Sidamon-Eristoff AE, Gilson A, Chartash D. We Role of Large Language Models in Medical Education: Applications and Implications. JMIR Med Educ. 2023 Aug 14;9:e50945.
8. KunzeKN,NwachukwuBU,CoteMP,RamkumarPN.LargeLanguageModelsAppliedtoHealthCare Tasks May Improve Clinical Efficiency, Value of Care Rendered, Research, and Medical Education. ArthroscJArthroscRelatSurgOffPublArthroscAssocNAmIntArthroscAssoc.2025Mar;41(3):547–56.
9. VrdoljakJ,BobanZ,VilovićM,KumrićM,BožićJ.AReviewofLargeLanguageModelsinMedical Education,ClinicalDecisionSupport,andHealthcareAdministration.HealthcBaselSwitz.2025Mar 10;13(6):603.
10. KimJ,VajraveluBN.AssessingtheCurrentLimitationsofLargeLanguageModelsinAdvancingHealth CareEducation.JMIRFormRes.2025Jan16;9:e51319.
11. AETCOM_book.pdf [Internet]. [cited 2022 Feb 12]. Available from: https:/www.nmc.org.in/wp-content/uploads/2020/01/AETCOM_book.pdf
12. RaviA,NeinsteinA,MurraySG.LargeLanguageModelsandMedical Education:PreparingforaRapid TransformationinHowTraineesWillLearntoBeDoctors.Sch.2023Sept;4(3):282–92.
13. JMIR Medical Education - Assessing AI Awareness and Identifying Essential Competencies: Insights From Key Stakeholders in Integrating AI Into Medical Education [Internet]. [cited 2025 Oct 24]. Available from: https: /mededu.jmir.org/2024/1/e58355/
14. SorteSR,RawekarAT,RathodSB,SuranaGandhiN.Future-readymedicine:AssessingtheneedforA.I. educationinIndianundergraduatemedicalcurriculum:Amixedmethodsurveyofstudentperspectives. JEducHealthPromot.2025May30;14:215.
15. von der Gracht HA. Consensus measurement in Delphi studies: Review and implications for future qualityassurance.TechnolForecastSocChange.2012Oct1;79(8):1525–36.
16. Humphrey-MurtoS,VarpioL,GonsalvesC,WoodTJ.Usingconsensusgroupmethodssuchas Delphi andNominalGroupinmedicaleducationresearch.MedTeach.2017Jan;39(1):14–9.
17. LeeYM,KimS,LeeYH,KimHS,SeoSW,KimH,etal.DefiningMedicalAICompetenciesforMedical SchoolGraduates:OutcomesofaDelphiSurveyandMedicalStudent/EducatorQuestionnaireofSouth Korean Medical Schools. Acad Med J Assoc Am Med Coll. 2024 May 1;99(5):524–33.
18. IndianCouncilofMedicalResearch.Nationalethicalguidelinesforbiomedicalandhealthresearch involvinghumanparticipants.2017NewDelhiIndianCouncilofMedicalResearchAvailablefrom: https:/www.icmr.nic.in/sites/default/files/guidelines/ICMR_Ethical_Guidelines_2017.pdf
19. MIR MM, MIR GM, RAINA NT, MIR SM, MIR SM, MISKEEN E, et al. Application of Artificial Intelligence in Medical Education: Current Scenario and Future Perspectives. J Adv Med Educ Prof. 2023 July;11(3):133–40.
20. We Digital Health Competencies in Medical Education Framework: An International Consensus Statement Based on a Delphi Study | Medical Education | JAMA Network Open | JAMA Network [Internet].[cited2025Oct24].Availablefrom: https:/jamanetwork.com/journals/jamanetworkopen/fullarticle/2829788#google_vignette
21. LargeLanguageModelsinHealthcareandMedicalApplications:AReview-PMC[Internet].[cited2025 Oct24].Availablefrom:https:/pmc.ncbi.nlm.nih.gov/articles/PMC12189880/
22. Meng X, Yan X, Zhang K, Liu D, Cui X, Yang Y, et al. We application of large language models in medicine:Ascopingreview.iScience.2024Apr23;27(5):109713.
23. LargeLanguageModelsinBiomedicalandHealthInformatics:AReviewwithBibliometricAnalysis- PubMed [Internet]. [cited 2025 Oct 24]. Available from: https: /pubmed.ncbi.nlm.nih.gov/39463859/

24. Saroha S. Artificial Intelligence in Medical Education: Promise, Pitfalls, and Practical Pathways. Adv Med Educ Pract. 2025 June 14;16:1039–46.
25. Rodger D, MannSP, EarpB, SavulescuJ, Bobier C, BlackshawBP. Generative AI inhealthcareeducation: How AI literacy gaps could compromise learning and patient safety. Nurse Educ Pract. 2025 Aug 1;87:104461.
26. Weidener L, Fischer M. Proposing a Principle-Based Approach for Teaching AI Ethics in Medical Education. JMIR Med Educ. 2024 Feb 9;10:e55368.
27. Masters K. Ethical use of Artificial Intelligence in Health Professions Education: AMEE Guide No. 158. MedTeach.2023June;45(6):574–84.
28. AlamF,LimMA,ZulkipliIN.IntegratingAIinmedicaleducation :embracingethicalusageandcritical understanding. Front Med. 2023 Oct 13;10:1279707.
29. AbujaberAA,NashwanAJ.Ethicalframeworkforartificialintelligenceinhealthcareresearch:Apathtointegrity.WorldJMethodol.2024Sept20;14(3):94071.
30. FrancoD'SouzaR,MathewM,MishraV,SurapaneniKM.Twelvetipsforaddressingethicalconcernsin the implementation of artificial intelligence in medical education. Med Educ Online. 2024 Dec 31;29(1):2330250.
31. Çalışkan SA, Demir K, Karaca O. Artificial intelligence in medical education curriculum: An e-Delphi study for competencies. PLoS ONE. 2022 July 21;17(7):e0271872.
32. Artificialintelligenceinmedicaleducationcurriculum:Ane-Delphistudyforcompetencies-PubMed [Internet]. [cited 2025 Oct 24]. Available from: https: /pubmed.ncbi.nlm.nih.gov/35862401/.