

A Competency Framework for Medical AI Education: Mixed Methods Study

CAI, Chang <<http://orcid.org/0009-0005-8056-0706>>, ZHU, Gaoxia <<http://orcid.org/0000-0003-4589-0775>>, ZHOU, Shang-Ming <<http://orcid.org/0000-0002-0719-9353>>, NG, Olivia <<http://orcid.org/0000-0001-7573-3725>>, DUELL, Jamie <<http://orcid.org/0000-0002-8837-7843>>, HO, Weng Kin <<http://orcid.org/0000-0002-4585-9340>>, CHEN, Daisy Minghui <<http://orcid.org/0009-0003-4608-8378>>, LEE, Bennett <<http://orcid.org/0000-0002-2709-1972>>, LI, Fang <<http://orcid.org/0009-0008-6436-8388>>, LIU, Siyuan <<http://orcid.org/0000-0002-5367-4709>>, SUDARSHAN, Vidya <<http://orcid.org/0000-0003-4494-8226>>, WANG, Li Rong <<http://orcid.org/0000-0003-3546-9354>>, CHOI, Chanwoo <<http://orcid.org/0009-0002-3032-9870>> and FAN, Xiuyi <<http://orcid.org/0000-0003-1223-9986>>

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/37472/>

This document is the Published Version [VoR]

Citation:

CAI, Chang, ZHU, Gaoxia, ZHOU, Shang-Ming, NG, Olivia, DUELL, Jamie, HO, Weng Kin, CHEN, Daisy Minghui, LEE, Bennett, LI, Fang, LIU, Siyuan, SUDARSHAN, Vidya, WANG, Li Rong, CHOI, Chanwoo and FAN, Xiuyi (2026). A Competency Framework for Medical AI Education: Mixed Methods Study. *JMIR Medical Education*, 12: e91116. [Article]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

Original Paper

A Competency Framework for Medical AI Education: Mixed Methods Study

Chang Cai^{1,2}, MAS; Gaoxia Zhu³, PhD; Shang-Ming Zhou⁴, PhD; Olivia Ng^{1,2}, PhD; Jamie Duell^{5,6}, PhD; Weng Kin Ho³, PhD; Daisy Minghui Chen¹, MAS; Bennett Lee^{1,7,8}, PhD; Fang Li⁹, PhD; Siyuan Liu⁹, PhD; Vidya Sudarshan⁹, PhD; Li Rong Wang^{9,10}, BSc; Chanwoo Choi⁹, MAS; Xiuyi Fan^{1,2,9,11}, PhD

¹Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

²National Health Group, Singapore, Singapore

³National Institute of Education, Nanyang Technological University, Singapore, Singapore

⁴Center for Health Technology, Faculty of Health, University of Plymouth, Plymouth, United Kingdom

⁵School of Computing and Digital Technologies, Sheffield Hallam University, Sheffield, United Kingdom

⁶Centre of Excellence in AI and Robotics, Sheffield Hallam University, Sheffield, United Kingdom

⁷Singapore Immunology Network (SIgN), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

⁸Infectious Disease Labs (ID Labs), Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

⁹College of Computing and Data Science, Nanyang Technological University, Singapore, Singapore

¹⁰Centre for Frontier AI Research, Agency for Science, Technology and Research (A*STAR), Singapore, Singapore

¹¹Centre of AI in Medicine, Nanyang Technological University, Singapore, Singapore

Corresponding Author:

Xiuyi Fan, PhD

Lee Kong Chian School of Medicine, Nanyang Technological University

50 Nanyang Avenue

Singapore

Singapore

Phone: 65 82634539

Email: xyfan@ntu.edu.sg

Abstract

Background: Although artificial intelligence (AI) is increasingly adopted in health care, clinicians face barriers, including insufficient understanding, limited trust, and challenges in interpreting AI outputs. Existing frameworks, such as the United Nations Educational, Scientific and Cultural Organization (UNESCO) AI competency framework, lack clinical specificity. Additionally, there remains limited evidence on framework-based training programs for medical professionals.

Objective: This study aimed to (1) develop a medical AI competency framework and (2) apply the framework to design an AI training program and pilot 1 module.

Methods: We conducted a mixed methods study comprising 2 phases. In Study 1, we developed a medical AI competency framework by integrating the UNESCO AI framework with the Miller pyramid model. The framework was refined through expert input from 24 stakeholders (6 hospital administrators, 8 medical professionals, and 10 university instructors). Expert responses were analyzed using deductive content analysis based on predefined codebooks. In study 2, we designed an AI training program based on the framework, and evaluated it using a 2-round Delphi process. Nine educators with expertise in instructional design, medical education, and AI participated in the Delphi study. Consensus in round 1 was defined as an IQR \leq 1, agreement score proportion $>75\%$, and full-score frequency $>49.23\%$; in round 2, consensus was defined as agreement score proportion $>80\%$. A pilot workshop with 28 participants and 4 instructors assessed the feasibility of 1 module using self-reported measures of satisfaction, engagement, and confidence.

Results: A 6D 4-level medical AI competency framework was developed. Among the 24 experts, 19 (79.17%) mentioned competencies related to AI foundations, and 23 (95.83%) mentioned competencies related to application skills. The framework was translated into a 5-module training program, covering patient-centered and ethical AI, privacy and security, medical AI applications, bias and health equity, and generative AI in health care. Each module included 5 elements: content, learning goals, teaching activities, learning resources, and assessment. The Delphi process achieved complete consensus across all 25 elements of the training program. The pilot workshop indicated high participant satisfaction (mean 4.00, SD 0.52), good engagement (mean 3.80, SD 0.71-mean 4.05, SD 0.51), and moderate self-reported confidence (mean 3.63, SD 0.53). These

findings suggest that the module was feasible, although outcomes should be interpreted cautiously given the self-reported and short-term nature of the evaluation.

Conclusions: The framework provides a structured reference for AI training program design in medical education. The workshop findings provide preliminary support for the feasibility of the program's technical module, while also highlighting the need for broader and longer-term evaluations. Future work should expand the framework and training program to new regions and delivery formats (eg, semester-long courses and continuing medical education) and evaluate their long-term impact.

JMIR Med Educ 2026;12:e91116; doi: [10.2196/91116](https://doi.org/10.2196/91116)

Keywords: artificial intelligence; AI; AI competency framework; AI literacy; AI training; medical education; medical professionals; United Nations Educational, Scientific and Cultural Organization; UNESCO

Introduction

Artificial intelligence (AI) is transforming health care by enabling early diagnosis of diseases, real-time health monitoring, and personalized health care [1-4]. It also contributes to health care management by enhancing patient satisfaction and optimizing hospital resources [5]. Nevertheless, clinicians encounter multiple barriers to AI adoption, including risks of misuse, insufficient understanding, limited trust in AI, and challenges in interpreting AI outputs [6,7].

These challenges underscore the need for clinicians to acquire AI competencies, as "AI is not meant to replace medical professionals, but the ones using AI will probably replace those who don't" [8, page 4]. Additionally, responsible AI adoption depends not only on external safeguards [9] but also on the capacities of medical professionals to interpret, evaluate, and oversee AI use in health care [10-16].

However, a dedicated AI competency framework tailored to medical professionals remains lacking. Global initiatives, such as the United Nations Educational, Scientific and Cultural Organization (UNESCO)'s AI competency framework [17], offer general guidelines for AI literacy but do not address the specific requirements of clinical practice. AI in health care requires unique considerations, including patient-centered decision-making, data governance, and ethical responsibility [11,12,18-20]. For example, clinicians need to explain AI-supported recommendations to patients, assess the reliability and limitations of AI outputs, and ensure that the use of AI aligns with patient values [10-12]. Moreover, proposed health care-related AI competencies are often limited to specific specialties, educational levels, or traditional AI technologies [13,15]. Emerging technologies, such as generative AI (GenAI), further highlight the insufficiency of existing proposed AI competencies. Although GenAI has considerable potential for health care [21], it also introduces challenges, including hallucinations and source uncertainty [22,23]. These challenges differ from those of traditional AI systems and require new competencies, such as verifying AI-generated content and ensuring clinician oversight [21,24]. However, few studies have examined the specific competencies required for medical professionals to use GenAI safely and effectively.

Furthermore, the absence of a dedicated competency framework has implications for medical education. Existing

medical AI education initiatives, such as extracurricular workshops, online courses, and specialized lectures [25-29], often focus on specific topics rather than a shared foundational competency structure. As a result, there is limited guidance on how AI skills should be developed, assessed, and integrated into curricula. Recent reviews have highlighted this gap, noting the lack of framework-guided curriculum design in medical AI education [30,31]. More broadly, AI education should not be limited to technical content coverage alone. It also requires structured pathways and guidance to help learners address application risks and support responsible implementation [32].

These gaps highlight the need for a comprehensive, clinically grounded, and future-oriented medical AI competency framework. Such a framework should be directly applicable to curriculum development. To address these issues, this study aims to answer the following research questions:

- What core AI competencies should medical professionals possess to meet the evolving demands of AI-enabled health care?
- How can we use an AI competency framework to build a training program?

This study makes 3 contributions. First, we develop a medical AI competency framework, aligning AI competencies with clinical demands. Second, we showed the practical application of the framework by designing an AI training program. Third, we provide preliminary evidence of feasibility through pilot implementation, offering insights into the design and delivery of AI curricula in medicine. This research may contribute to the integration of AI into medical education and practice.

A preliminary version of this work was presented in the Artificial Intelligence in Education 2025 conference proceedings [33]. That paper reported on the development and refinement of the medical AI competency framework. This study substantially extends that foundational work in several important ways. First, it provides a more detailed and rigorous description of the framework development and expert-informed refinement process. Second, it extends the framework by incorporating observable behaviors, making it more actionable for teaching and assessment. Third, building on the framework, a training program for medical AI education was developed and evaluated through a 2-round Delphi process. Finally, a pilot study of 1 module

of the program was implemented, and student feedback was collected to examine its feasibility. Overall, this study extends the preliminary work by providing a more actionable framework, showing its application in curriculum development, and offering preliminary evidence of implementation feasibility through a pilot module.

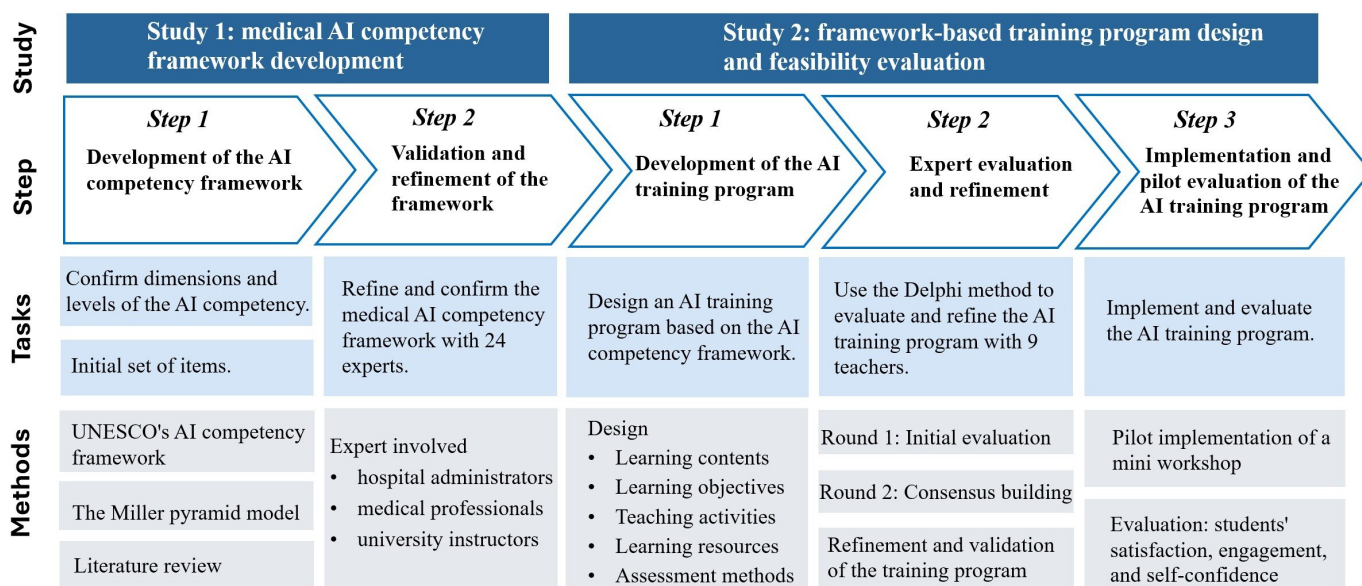
Methods

Overview

This research consists of 2 studies, as shown in Figure 1. Study 1 focused on the development of the medical AI

competency framework. An initial version of study 1 was reported in our Artificial Intelligence in Education 2025 conference paper [33]; in this paper, this component is described in greater methodological detail, while study 2 is newly reported. Study 2 showed one practical application of this framework. We first designed a training program grounded in the competency framework. The program was subsequently evaluated and refined through a 2-round Delphi process. Finally, a mini workshop was carried out to examine the program’s feasibility.

Figure 1. Five-step development of the medical artificial intelligence (AI) competency framework and training program. UNESCO: United Nations Educational, Scientific and Cultural Organization.



Study 1: Development of the Medical AI Competency Framework

Step 1: Initial Medical AI Competency Framework Conceptualization

This step establishes the theoretical foundation for the initial medical AI competency framework. The framework was developed based on (1) UNESCO’s AI competency framework [17], (2) the Miller pyramid [34], and (3) prior literature on AI competencies in medicine [11,12,15,35,36].

Theoretical Foundation

To conceptualize medical AI competence, we first drew on UNESCO’s AI competency framework [17], a general guideline for AI literacy development. It outlines four core domains [17]: (1) human-centered mindset (developing a human-focused perspective on AI’s benefits, risks, and proportionality in addressing societal and environmental needs), (2) ethics of AI (acquiring social and ethical competencies to guide the responsible use of AI throughout its lifecycle), (3) AI techniques and applications (combining conceptual understanding with practical skills through the real-world application of AI tools), and (4) AI system design

(developing technical proficiency in designing, implementing, and refining AI systems).

Building on these domains and prior literature [11,12,15,17,35,36], we derived 6 competency dimensions tailored to the needs of health care practice. From UNESCO’s *human-centered mindset*, we introduced *patient-centered AI in health care* to capture clinicians’ ability to interpret and apply AI-supported decisions while preserving autonomy and maintaining human oversight. Extending UNESCO’s *ethics of AI*, we proposed *ethics and transparency in clinical AI*, highlighting ethical principles in medical AI implementation. From UNESCO’s *AI techniques and applications*, we derived *technical proficiency in medical AI applications*, emphasizing practical AI use by clinicians. Integrating UNESCO’s *AI system design* with a focus on health care equity [37,38], we introduced *bias mitigation and health equity in AI design*. Considering the importance of data privacy and governance in medicine [19,20], we incorporated *data privacy, security, and compliance in health care AI*. Finally, recognizing the expanding role and challenges of GenAI in health care [21-23], we proposed *GenAI for health care*.

Proficiency Levels

Rather than UNESCO's 3-level progression (understand, apply, and create), we adopted the Miller pyramid model [34] to guide competency development. Unlike generic progression models, the Miller pyramid provides a clinically oriented structure that differentiates knowledge, application, simulated performance, and real clinical practice [34], making it well-suited for medical professionals. Widely used in medical education to assess clinical skills [39,40], it defines step-wise progression for competency development. The Miller pyramid includes the following levels [34]:

1. Know (knowledge): acquisition of foundational knowledge.
2. Know how (competency): the ability to apply knowledge in context.
3. Show how (performance): demonstration of skills in simulated settings.
4. Do (action): application of skills in real clinical practice.

These 6 dimensions and 4 levels jointly constituted the initial medical AI competency framework (Table S1 in [Multimedia Appendix 1](#)).

Table 1. Demographics and selection criteria of framework validation experts.

Experts	Number of experts	Department	Selection criteria
Hospital administrators	6	Hospitals and government agencies	Leaders from hospitals or government agencies with more than 5 years of experience in policy-making and management.
University instructors	10	A university and health care organization	Teachers from universities or health care organizations with more than 5 years of teaching experience in AI ^a or medicine.
Medical professionals	8	Hospitals	Medical professionals with more than 5 years of clinical experience and prior exposure to AI applications.

^aAI: artificial intelligence.

Data Collection

Experts were invited via email between August 2024 and December 2024 and completed an anonymous online survey with open-ended questions. Responses were collected anonymously to ensure participant confidentiality. Participants had the opportunity to review and modify their responses before submission.

The survey included the following questions for all experts:

- What AI-related skills do you think medical professionals should acquire?
- What topics in AI for medicine, do you consider essential for medical professionals to learn?
- For each AI-related skill mentioned above, to what extent do you think medical professionals should possess it (eg, basic awareness, practical knowledge, high-level skills, or advanced expertise)?

For medical professionals, additional questions included the following:

Step 2: Expert-Informed Framework Refinement

In this step, the initial framework was used to build deductive codebooks (Tables S2-S7 in [Multimedia Appendix 1](#)). The codebook was used to analyze experts' responses to the open-ended questions.

Expert Panel

To examine whether the initial framework aligned with expert views and to refine it where needed, a total of 24 experts from Singapore were selected, including 6 hospital administrators, 8 medical professionals, and 10 university instructors (Table 1). These diverse stakeholders ensured alignment of the framework with health care system needs and pedagogical principles. The selection of Singaporean experts was due to logistical constraints, and future studies will expand to include international perspectives.

- Based on your own work experience, how can AI complement medical practices?
- What are the main barriers you have encountered when using or attempting to use AI?
- What challenges have you faced, or do you anticipate facing, when co-designing and co-developing a medical AI product?

Data Analysis

Expert feedback was analyzed through deductive content analysis using predefined codebooks (Tables S2-S7 in [Multimedia Appendix 1](#)). The codebooks were developed based on the initial medical AI competency framework (Table S1 in [Multimedia Appendix 1](#)), which was derived from our prior work [33]. Two researchers independently coded one-third of the responses using the codebooks to assess inter-rater reliability. The κ coefficient was 0.81, indicating strong agreement [41]. Researchers assigned each response to the most relevant competency category. Comments that did not fit existing definitions or suggested a new competency were flagged, discussed iteratively, and refined. In this study, we did not identify expert comments that justified adding

a new dimension. However, several code definitions were refined to better reflect the expert comments while preserving the broader 6D structure. Discrepancies were resolved through iterative discussion until full consensus was reached. After establishing reliability, 1 researcher coded the remaining responses using the codebooks.

Study 2: Framework-Based Training Program Design and Feasibility Evaluation

Step 1: Framework-Based AI Training Program Development

We developed an AI training program as an application of the framework. Prior AI in Medicine initiatives suggested that combining lectures with case-based learning, hands-on activities, and problem-based learning supports the application of AI concepts in clinical contexts and enhances learner engagement [25,26,42,43]. Additionally, flipped classrooms have been extensively adopted to improve learning outcomes [25,44]. Integration of online and offline resources can support teaching design by aligning preclass learning, in-class interaction, and practice [45]. To support experiential and reflective learning, the Kolb experiential learning cycle and the Gibbs reflective cycle were introduced [46,47]. These pedagogical and theoretical principles informed the design of the program's learning activities and assessment strategies.

The program design included 5 core components: content, learning objectives, teaching activities, learning resources, and assessment methods. Modules were aligned with the competency framework, and learning objectives were mapped to the Miller pyramid. We also adopted teaching activities

aligned with the Miller pyramid. Lectures build foundational knowledge (know); case studies, discussions, and hands-on activities enable application and demonstration (know how and show how); and problem-based learning fosters real-world problem-solving (do). Various resources were provided: (1) curated datasets for hands-on data analysis, (2) case studies on AI's clinical application, and (3) online platforms for coding.

Step 2: Evaluation and Refinement of the Training Program

To validate the AI training program, we used the Delphi method, inviting 9 educators with expertise in AI, medical education, and instructional design.

Course Evaluation Panel

The panel was selected based on three criteria: (1) participants have designed at least 1 course, (2) participants have teaching experience in higher education, and (3) participants are from the fields of education, medicine, or AI.

As detailed in Table 2, the course evaluation panel included 2 education specialists from the education department of a public university in Singapore, 3 AI teachers from the computer science department of a public university in Singapore, and 4 medical teachers from the medical departments of public universities in Singapore and the United Kingdom. This multidisciplinary composition ensured a balanced integration of pedagogical, AI, and clinical expertise, providing a comprehensive evaluation of the training program.

Table 2. Demographics of the course evaluation panel participant.

Participant	Number of experts	Department	Nationality
Education experts	2	Educational department	Singapore
AI ^a educators	3	Computer science department	Singapore
Medicine educators	4	Medicine department	Singapore and the United Kingdom

^aAI: artificial intelligence.

Delphi Method

The Delphi method is a well-established, anonymous group process designed to achieve consensus among participants [48,49]. This method has been widely used in previous studies to support and validate course design [50-52]. We conducted a 2-round Delphi process over 1 month to reach consensus on the AI training program design. No consultation on methods took place. Participants were invited by email and completed online questionnaires. The process followed a quasi-anonymous design: panelists were not informed of each other's identities, and no direct interaction occurred between participants. Responses were deidentified prior to analysis.

In round 1, participants reviewed the proposed 5-module training program and rated 25 items (covering content, learning objectives, teaching activities, learning resources, and assessment) using a 5-point Likert scale (1="strongly

disagree" to 5="strongly agree"). Open-ended comments were also collected to inform revisions. Qualitative feedback from round 1 was synthesized by 1 researcher by deduplicating and grouping similar suggestions. The grouped suggestions were then summarized into specific revisions (eg, module, element, and corresponding suggestions). A second researcher reviewed the synthesized feedback, and discrepancies were resolved through discussion. The aggregated feedback was then provided to the same panel in round 2. Participants re-evaluated the revised items based on the summarized feedback from round 1 using the same rating scale. The round 1 and round 2 surveys can be accessed in [Multimedia Appendix 2](#).

Data Analysis and Consensus

Panel ratings were analyzed using the IQR, agreement score (AS) proportion (proportion of ratings ≥ 4), and full-score

(FS) frequency (proportion of ratings=5), consistent with previous studies [52-57]. All consensus thresholds were prespecified. In round 1, an item was considered to have reached consensus if it met all 3 criteria: IQR ≤ 1 , AS $>75\%$, and FS above a predefined threshold. Following previous work [58], the threshold for FS was calculated as mean-SD, which yielded 49.23% (63.11%-13.88%). In this study, FS $>49.23\%$ corresponds to more than half of the panel assigning the highest score. Items that did not meet these criteria were carried forward to round 2. In round 2, consensus was defined as achieving AS $>80\%$ [55-57].

Step 3: Pilot Implementation of a Mini Workshop

To assess the feasibility of the designed AI training program, we conducted a pilot mini workshop. Due to time constraints, the workshop focused primarily on module 3 (medical AI applications), while the remaining modules remained at the design stage and were not fully implemented. Module 3 was selected because it represented the core technical component of the program. A 1-day in-person workshop was held in July 2025. A total of 28 participants and 4 instructors attended. During in-class sessions, participants worked in small groups to analyze real-world AI use cases in health care. The

teaching activities included lectures, hands-on sessions, group presentations, and reflective discussions.

After the workshop, participants were invited to complete a voluntary and anonymous online postclass survey (Table 3). The 30 items were adapted from validated instruments: the Student Satisfaction and Self-Confidence in Learning scale and the Student Engagement scale [59, 60], covering satisfaction, self-confidence, and engagement in learning. We also collected instructors' insights through a survey with 6 open-ended questions. Full surveys are available in [Multimedia Appendix 2](#). Participants were able to review and modify their responses before submission. Descriptive statistics were used to summarize questionnaire responses. Open-ended responses were analyzed using the 6-step thematic analysis approach, developed by Braun and Clarke [61]. One researcher read all responses repeatedly and generated initial codes. These codes were then grouped into preliminary themes. A second researcher subsequently reviewed the coding and the preliminary thematic structure. Any discrepancies in code interpretation or theme grouping were discussed between the 2 researchers until consensus was reached. Through this process, the final themes were named and synthesized.

Table 3. Structure of the postworkshop survey.

Participants, dimension, and subdimension	Number of items	Type
Students		
Satisfaction	5	5-point Likert scale
Engagement		
Behavioral engagement	5	5-point Likert scale
Emotional engagement	4	5-point Likert scale
Cognitive engagement	8	5-point Likert scale
Self-confidence	8	5-point Likert scale
Overall feedback	2	Open-ended questions
Instructors		
Overall feedback	6	Open-ended questions

Ethical Considerations

The study was reviewed and approved by the institutional review board of Nanyang Technological University (IRB-2024-928). Written informed consent was waived by the review board, as the study posed minimal risk and involved anonymous surveys conducted online. Online informed consent was provided to all participants prior to data collection, and participation was entirely voluntary. All survey data were collected anonymously with no personally identifiable information recorded, and access was restricted to the research team. Participants did not receive any compensation.

Results

Results of Study 1: Medical AI Competency Framework Development

Expert Validation and Framework Refinement

Most experts highlighted the importance of "AI foundations" (n=19) and "application skills" (n=23). These findings highlight the need for core AI knowledge and the ability to learn, apply, and co-develop AI solutions within medical workflows. As a medical professional emphasized, "gaining basic AI knowledge, gaining data analysis skills, and applying AI in clinical settings to improve patient care" were seen as essential competencies. Similarly, a hospital administrator noted that expected competencies include being "able to identify problems that can be solved by AI, understanding the process of deployment and evaluation of AI projects,

applying AI knowledge in innovation, and possessing data analysis skills.”

In addition to technical competencies, 16 experts highlighted the importance of “critical reflections on AI,” emphasizing the ethical and safe use of AI in medicine. As a medical professional pointed out, “AI ethics concepts are similar to medical ethics concepts ... [professionals should] learn more about AI ethics and governance.” Moreover, 5 experts emphasized the integration of GenAI into clinical workflows. For example, a medical professional stated, “Using ChatGPT to conceptualize brief background of certain issues.” Based on expert feedback for greater clarity, “application skills” and “co-design equitable AI systems” were redefined to emphasize collaboration with engineers. For example, 1 hospital administrator highlighted:

Not to become a coder/developer/engineer, but to get skillset to be able to comprehensively lead teams and

run AI projects. Some technical + system thinking + appreciation of broader landscape.

Furthermore, the definition of “real-world compliance” was expanded to stress the proper handling of both sensitive and nonsensitive data in clinical contexts. Detailed results are provided in Table S8 in [Multimedia Appendix 3](#).

Final Medical AI Competency Framework

Table 4 and Table 5 show the final framework, which comprises 6 dimensions across 4 levels, ranging from foundational knowledge (know) to clinical practice (do).

Table 4. Medical artificial intelligence (AI) competency framework (part 1; dimensions 1-3).

Aspects	Progression			
	Know	Know how	Show how	Do
Patient-centered AI ^a in health care	<ul style="list-style-type: none"> Basic understanding: understand the foundational impact of AI on patient rights and autonomy Observable behaviors: identification of patient rights affected by AI use; recognition of patient autonomy concerns in AI-supported care; identification of patient-centered risks in AI use 	<ul style="list-style-type: none"> Human advancement: learn how to apply AI to enhance treatment planning, automate care processes, and personalize health care for individual patients Observable behaviors: explanation of how AI supports treatment and personalization; identification of appropriate AI use in care processes 	<ul style="list-style-type: none"> Patient-centered decision-making: use AI effectively to support decision-making, ensuring patient well-being is prioritized Observable behaviors: interpretation of AI outputs; comparison of AI recommendations with patient needs; justification of AI decisions based on patient well-being 	<ul style="list-style-type: none"> Patient-centered practice: integrate AI tools into clinical practice, taking responsibility for incorporating them into patient care workflows to improve care quality and outcomes Observable behaviors: communication with patients about AI outputs in routine care workflows; integration of AI into patient care workflows with clinician oversight; accountable use of AI to support care quality
Ethics and transparency in clinical AI	<ul style="list-style-type: none"> Critical reflections on AI: comprehend the basic ethical principles guiding the use of AI in health care settings Observable behaviors: identification of core ethical principles in clinical AI; recognition of the importance of transparency in health care AI; recognition of ethical concerns in AI-supported care 	<ul style="list-style-type: none"> Safe and responsible use: know how to use AI in ways that ensure ethical considerations and transparency Observable behaviors: explanation of ethical principles in health care AI use; consideration of transparency when planning AI-supported decisions; identification of ethical risks in planned AI use 	<ul style="list-style-type: none"> Ethical decision-making: demonstrate ethical considerations during AI-enabled decision-making in simulated settings Observable behaviors: appraisal of ethical issues in simulated AI-supported cases; justification of decisions based on ethical and transparency considerations 	<ul style="list-style-type: none"> Ethics by design: implement ethical standards in real-world AI applications, evaluating and improving systems for ethical compliance Observable behaviors: adherence to ethical standards in clinical AI practice; monitoring of AI use for ethical compliance; improvement of AI systems for ethical and transparent practice; documentation of ethical concerns in practice
Data privacy, security, and	<ul style="list-style-type: none"> Privacy awareness: remember regulations 	<ul style="list-style-type: none"> Regulatory adherence: know how to handle 	<ul style="list-style-type: none"> Secure data practices: manage patient data 	<ul style="list-style-type: none"> Real-world compliance: process clinical data under

Aspects	Progression			
	Know	Know how	Show how	Do
compliance in health care AI	and the importance of safeguarding patient data <ul style="list-style-type: none"> Observable behaviors: awareness of privacy and data governance regulations; recognition of the importance of patient data protection 	patient data securely, using techniques like anonymization and governance <ul style="list-style-type: none"> Observable behaviors: explanation of secure data handling practices; identification of data anonymization approaches 	in simulated research projects, adhering to privacy laws <ul style="list-style-type: none"> Observable behaviors: application of data protection measures in structured tasks; management of data under privacy constraints in projects 	specific conditions, ensuring all data handling and AI applications comply with privacy and security standards in clinical and research settings <ul style="list-style-type: none"> Observable behaviors: secure handling of clinical data in routine workflows; responsible data use in clinical and research settings

^aAI: artificial intelligence.

Table 5. Medical artificial intelligence (AI) competency framework (part 2; dimensions 4-6).

Aspects	Progression			
	Know	Know how	Show how	Do
Technical proficiency in medical AI ^a applications	<ul style="list-style-type: none"> AI foundations: acquire foundational knowledge of AI concepts and techniques Observable behaviors: awareness of core AI concepts; recognition of basic AI methods and models; understanding of algorithms in health care AI 	<ul style="list-style-type: none"> Health care AI implementation: learn how to apply AI methods effectively in medical contexts Observable behaviors: explanation of AI applications in medical contexts; explanation of AI support for clinical decision-making; identification of appropriate clinical use cases for AI 	<ul style="list-style-type: none"> Application skills: grasp the ability to perform and cocreate basic AI-related programming or data analysis tasks Observable behaviors: performance of basic AI-related programming tasks; performance of basic data analysis tasks 	<ul style="list-style-type: none"> Evaluate and create AI: critically assess existing AI systems and develop solutions tailored to specific health care needs Observable behaviors: evaluation of AI systems for clinical relevance; development of AI solutions for health care needs; implementation of AI tools in real-world settings
Bias mitigation and health equity in AI design	<ul style="list-style-type: none"> Foundational understanding of bias: Identify how AI impacts health equity and its potential to either mitigate or exacerbate disparities Observable behaviors: awareness of bias in health care AI; recognition of AI-related health disparities; understanding of AI's impact on health equity 	<ul style="list-style-type: none"> Bias detection and analysis: know how to identify, evaluate, and address bias in AI systems within health care contexts Observable behaviors: explanation of bias sources in health care AI; evaluation of bias risks in AI use; identification of bias risks in data and AI systems 	<ul style="list-style-type: none"> Equity-oriented practices: select and apply AI tools that minimize bias and enhance fairness, ensuring health equity in controlled design or assessment tasks Observable behaviors: selection of AI tools with fairness considerations; application of bias mitigation approaches in structured tasks 	<ul style="list-style-type: none"> Co-design equitable AI systems: cocreate and refine AI solutions to enhance health equity, using iterative feedback to address systemic disparities Observable behaviors: co-design of AI solutions for health equity; refinement of AI systems to reduce bias; iterative improvement of AI for equitable practice
GenAI ^b for health care	<ul style="list-style-type: none"> GenAI basics: identify the basic 	<ul style="list-style-type: none"> Interpretative proficiency: know 	<ul style="list-style-type: none"> Skillful development: use GenAI, 	<ul style="list-style-type: none"> Integration in practice: independently integrates,

Aspects	Progression			
	Know	Know how	Show how	Do
	principles, limitations, and benefits of GenAI <ul style="list-style-type: none"> • Observable behaviors: identification of core GenAI concepts; understanding of potential clinical uses of GenAI; recognition of GenAI limitations in health care 	how to interpret AI outputs, adapt data, and identify risks like hallucinations. <ul style="list-style-type: none"> • Observable behaviors: Explanation of strategies for using GenAI tools; identification of hallucination and reliability risks 	validate outputs, and communicate effectively in clinical scenarios <ul style="list-style-type: none"> • Observable behaviors: validation of GenAI outputs in structured clinical tasks; writing of safe prompts; documentation of output limitations and need for human review 	evaluates, and oversees GenAI in real clinical workflows <ul style="list-style-type: none"> • Observable behaviors: Integration of GenAI into real clinical workflows; verification of GenAI outputs before use in patient care; disclosure of limitations and maintenance of clinician oversight

^aAI: artificial intelligence.

^bGenAI: generative artificial intelligence.

Translation of the Study 1 Findings Into the Study 2 Training Program

The final competency framework developed in study 1 served as the reference for the training program design in study 2. Specifically, 6 competency dimensions informed the learning topics and content of the program, while 4 Miller-aligned levels guided the design of learning objectives and assessment tasks. Learning activities were further developed in alignment with both the competency framework and

educational theories. As shown in [Table 6](#), the findings from study 1 were translated into the training program by mapping 6 dimensions of the medical AI competency framework into 5 modules. “Patient-centered AI in health care” and “ethics and transparency in clinical AI” were combined into a single module because of their close conceptual relationship in clinical practice. In this way, study 2 represents one practical application of the competency framework developed in study 1.

Table 6. Mapping of the medical artificial intelligence (AI) competency framework to the AI training program.

Module	Competency dimension	Learning focus
Module 1: patient-centered and ethical AI in Health Care	<ul style="list-style-type: none"> • Patient-centered AI • Ethics and transparency in clinical AI 	Ethical principles of AI use in health care, patient autonomy, fairness, transparency, and responsible clinical decision-making
Module 2: AI privacy and security in health care	<ul style="list-style-type: none"> • Data privacy, security, and compliance in health care AI 	Data protection, regulatory frameworks (eg, HIPAA ^a and GDPR ^b), and secure handling of health care data in AI applications
Module 3: medical AI applications	<ul style="list-style-type: none"> • Technical proficiency in medical AI applications 	Fundamental AI concepts, model implementation, and interpretation of AI outputs in health care contexts
Module 4: AI bias and health equity	<ul style="list-style-type: none"> • Bias mitigation and health equity in AI design 	Identification of bias in AI systems, fairness considerations, and strategies to promote equitable health care outcomes
Module 5: generative AI in Health Care	<ul style="list-style-type: none"> • Generative AI for health care 	Safe use of generative AI, including prompt design, output evaluation, and responsible integration into clinical workflows

^aHIPAA: Health Insurance Portability and Accountability Act.

^bGDPR: General Data Protection Regulation.

Results of Study 2: Framework-Based Training Program Design and Feasibility Evaluation

Delphi Round 1 and Round 2 Results

Nine experts participated in both Delphi rounds. After round 1, consensus was reached on 16 of the 25 (64%) training

program elements. The remaining 9 elements were carried forward to round 2 because they did not meet at least one of the predefined round 1 consensus criteria. In round 2, agreement was reached on the remaining 9 elements. [Table 7](#) shows the results of the 2-round survey. [Table S9](#) in [Multimedia Appendix 3](#) shows the full results of the 2-round survey.

Table 7. Results of the 2-round Delphi study.

Element of program and module	Round 1 ^a			Round 2 ^b
	Median (IQR)	AS ^c (%)	FS ^d (%)	AS (%)
Teaching activities				
Module 3	4 (1)	88.89	44.44	100
Learning resources				
Module 1	5 (2)	66.67	66.67	100
Module 2	5 (2)	66.67	66.67	88.89
Module 5	5 (2)	66.67	55.56	88.89
Learning assessments				
Module 1	4 (1)	88.89	44.44	100
Module 2	4 (1)	88.89	44.44	100
Module 3	4 (1)	88.89	33.33	100
Module 4	4 (1)	88.89	44.44	100
Module 5	4 (1)	88.89	44.44	100

^aIn round 1, consensus was defined as meeting all the 3 criteria: IQR \leq 1, AS >75%, and FS >49.23%.

^bIn round 2, consensus was defined as achieving AS>80%.

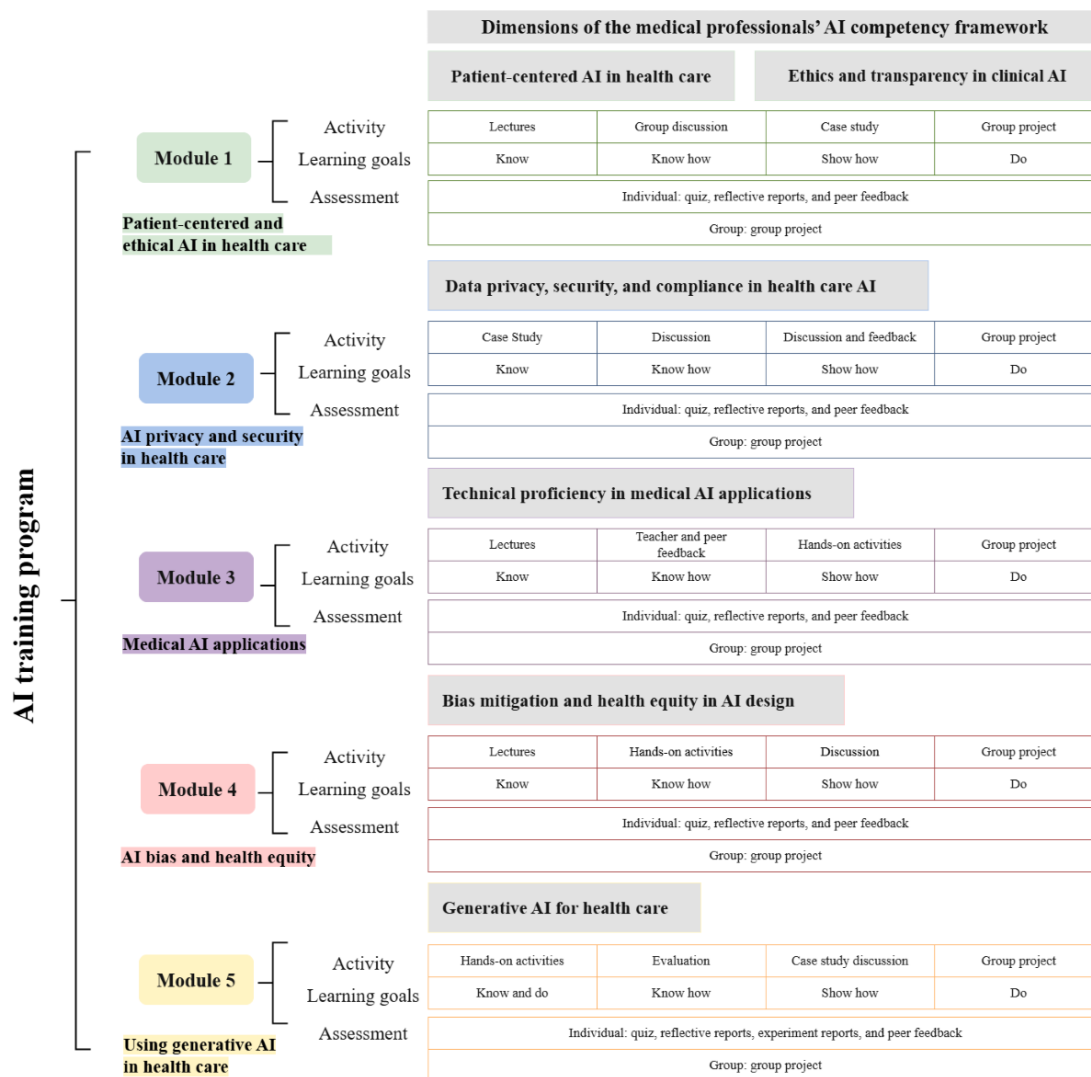
^cAS: agreement score proportion, the proportion of responses within a defined score range (eg, 4 and 5 on a 5-point scale) [53].

^dFS: full score frequency, the proportion of the highest scores [54].

Final Framework–Based AI Training Program

Figure 2 shows the overview structure of the 5-module AI training program. Multimedia Appendix 4 shows the final program design.

Figure 2. Structure of the artificial intelligence (AI) training program.



Module 1: Patient-Centered and Ethical AI in Health Care

This module aligns with the framework domains of “patient-centered AI in health care” and “ethics and transparency in clinical AI.” It introduces the ethical use of AI in health care, with a focus on patient-centered care, the impact of AI on

patient rights, and the principles of transparency in clinical AI applications. Medical professionals are guided to balance ethical decision-making with AI integration, ensuring patient autonomy and the protection of patient rights. For example, Table 8 shows example learning tasks and assessments aligned with the “show how” and “do” levels of the Miller pyramid.

Table 8. Example alignment of learning tasks and assessment methods.

Module and level	Tasks	Assessment methods
Module 1		
Show how	Analyze a health care case involving AI ^a -supported diagnosis; identify issues related to patient autonomy, fairness, transparency, and human oversight; justify how patient interests should be protected.	Case analysis report or presentation with rubric-based assessment of ethical analysis, patient-centered health care, and justification of recommendations.
Do	Propose a proposal for an AI tool in health care that respects patient rights, supports fair decision-making, and includes clear human oversight in clinical use.	Group project with rubric-based assessment of patient-centered design, ethical safeguards, transparency, and practical integration into health care workflow.
Module 2		
Show how	Analyze privacy and security case studies involving improper disclosure, disposal, or handling of patient data.	Case-based task with structured rubric assessing identification of privacy and security risks.

Module and level	Tasks	Assessment methods
Do	Design a health care AI project involving patient data; classify sensitive and nonsensitive data and propose protection measures.	Group project with rubric-based assessment of appropriateness of privacy and security measures and alignment with health care data governance requirements.
Module 3		
Show how	Run a preconfigured AI model using health care data; interpret performance metrics such as accuracy, F_1 -scores, precision, and recall; modify code or model settings (eg, hyperparameters).	Hands-on coding task with rubric-based assessment of code execution, interpretation of model outputs, and appropriateness of modifications.
Do	Adapt a basic AI model to a health care problem and present how the model could support clinical tasks, including its strengths and limitations.	Group project with rubric-based assessment of technical implementation, performance improvement, relevance to the health care problem, and critical interpretation of model limitations.
Module 4		
Show how	Use crafted imbalanced datasets to identify bias in model development; examine biased predictions using fairness metrics such as false positive rates.	Written reports assessed with structured rubric (bias detection, fairness analysis, and model interpretation).
Do	Design an AI model using a crafted dataset, apply bias mitigation methods such as oversampling, and evaluate whether the revised model improves fairness and equity.	Group project with rubric-based assessment of model design, bias mitigation, and fairness outcomes.
Module 5		
Show how	Write safe and effective prompts for clinical tasks; identify hallucinated or unsupported claims; analyze cases in which GenAI ^b produces incorrect medical advice; document limitations explicitly.	Case-based task with a structured rubric assessing prompt quality, error detection, justification of validation steps, and documentation of limitations; individual report evaluating AI-generated outputs.
Do	Design a GenAI-supported solution based on a health care task (eg, documentation or literature synthesis); implement validation steps; specify when human review is required.	Group project with rubric-based assessment of use-case appropriateness, output reliability, validation strategy, transparency of limitations, and workflow integration.

^aAI: artificial intelligence.

^bGenAI: generative artificial intelligence.

Module 2: AI Privacy and Security in Health Care

This module corresponds to the framework domain of “data privacy, security, and compliance in health care AI.” It covers knowledge of data privacy, security, and regulatory compliance in medical AI applications. Medical professionals will learn key regulations (eg, General Data Protection Regulation and Health Insurance Portability and Accountability Act [HIPAA]) and ensure compliance in clinical settings. For example, [Table 8](#) presents example learning tasks and assessments aligned with the “show how” and “do” levels of the Miller pyramid.

Module 3: Medical AI Applications

Aligned with the framework domain of “technical proficiency in medical AI applications,” this module provides knowledge of AI applications in health care. Medical professionals will learn the foundational principles of medical AI, explore AI applications, and gain hands-on experience in building and evaluating AI models. The module emphasizes application skills as well as the ability to evaluate and create clinically meaningful AI solutions. For example, [Table 8](#) shows example learning activities and assessments aligned with the “show how” and “do” levels of the Miller pyramid.

Module 4: AI Bias and Health Equity

This module corresponds to the framework domain of “bias mitigation and health equity in AI design.” It develops medical professionals’ foundational understanding of

algorithmic bias and its unequal impact on diverse patient populations. Through hands-on work with real datasets and explainable AI tools, they will examine how bias arises in data and affects models. The module further emphasizes bias detection, equity-oriented practices, and the co-design of fair and inclusive AI systems. For example, [Table 8](#) shows example learning activities and assessments aligned with the “show how” and “do” levels of the Miller pyramid.

Module 5: Using Generative AI in Health Care

Aligned with the framework domain of “generative AI for health care,” this module introduces medical professionals to the principles and clinical applications of GenAI. They will acquire knowledge of generative models, learn prompt engineering, and gain practical experience applying GenAI. The module emphasizes the safe, effective, and responsible integration of GenAI into clinical practice. This emphasis is particularly important in health-related large language model applications, where outputs may be unverifiable or clinically incorrect, highlighting the need for validation, source grounding, and clinician oversight [21,62]. For example, [Table 8](#) presents example learning activities and assessments aligned with the “show how” and “do” levels of the Miller pyramid.

Pilot Implementation (Mini Workshop) Results

Overview

Of the 28 participants, 15 reported their learning experiences ([Table 9](#)). All questionnaires had high internal consistency, with Cronbach α values greater than 0.8. The mean scores

were 4 (SD 0.52) for satisfaction, 3.8 (SD 0.71) for behavioral engagement, 4.05 (SD 0.51) for emotional engagement, 3.96 (SD 0.57) for cognitive engagement, and 3.63 (SD 0.53) for

self-confidence (Table 10). Figure 3 illustrates the distribution of participant responses across survey dimensions.

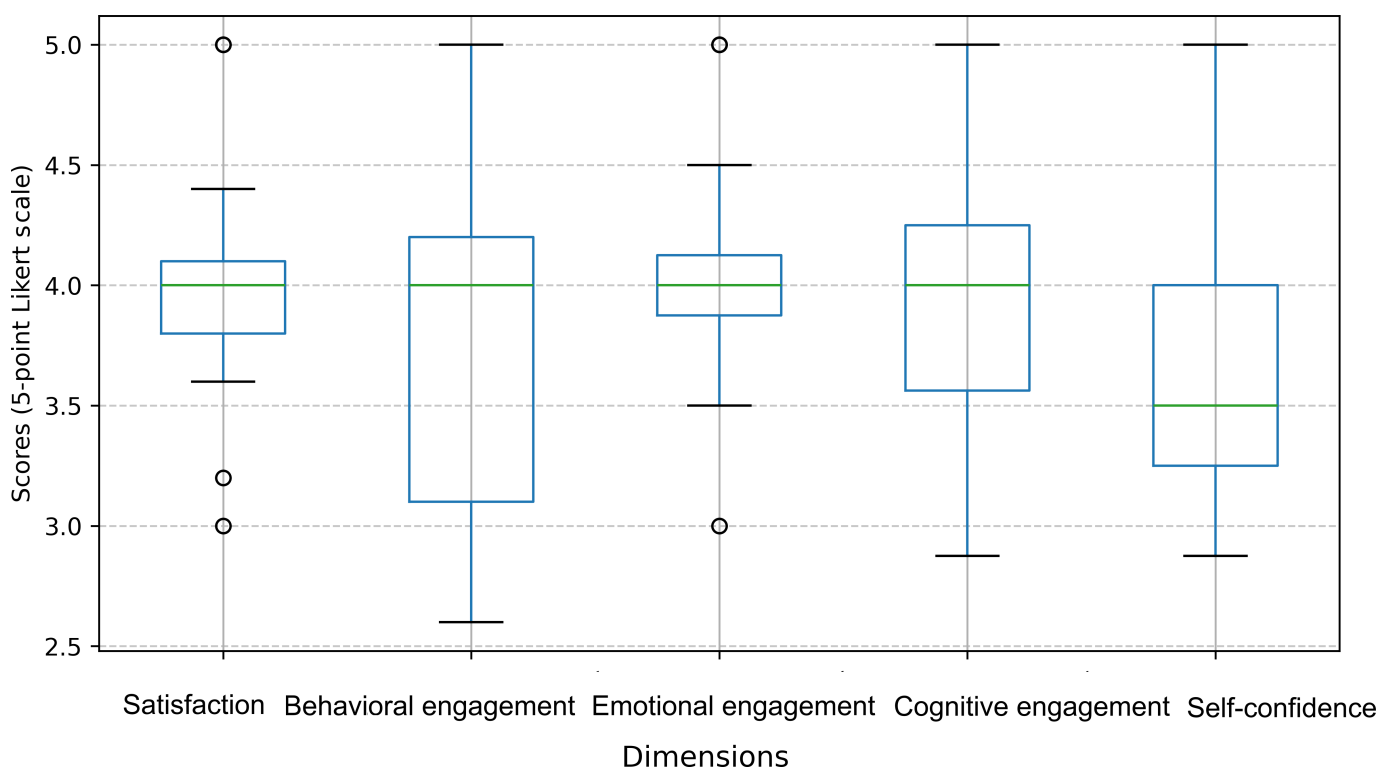
Table 9. Participant descriptions and demographics.

Characteristics and categories	Participants (n=28), n (%)	Survey respondents (n=15), n (%)
Roles		
Year 2 medical student	12 (42.9)	6 (40)
Year 3 medical student	4 (14.3)	2 (13.3)
Year 4 medical student	7 (25)	4 (26.7)
Year 5 medical student	5 (17.9)	3 (20)
Training stage		
Preclinical (year 1-2)	12 (42.9)	6 (40)
Clinical (year 3-5)	16 (57.1)	9 (60)
Motivation for attending		
Voluntary	28 (100)	15 (100)
Others	0 (0)	0 (0)

Table 10. Descriptive statistics of post-workshop survey dimension (n=15).

Dimension	Mean (SD)	Median (IQR)	Min-max
Satisfaction	4 (0.52)	4 (5-3)	3.00-5.00
Behavioral engagement	3.80 (0.71)	4 (5-2.60)	2.60-5.00
Emotional engagement	4.05 (0.51)	4 (5-3)	3.00-5.00
Cognitive engagement	3.96 (0.57)	4 (5-2.88)	2.88-5.00
Self-confidence	3.63 (0.53)	3.50 (5-2.88)	2.88-5.00

Figure 3. Distribution of 15 participant responses across the post-workshop survey dimension.



Qualitative findings from open-ended responses revealed 3 main themes that reflected participants' learning experiences and perceptions of the workshop.

Perceived Understanding of Coding and AI-Related Concepts

Participants reported learning coding basics, AI applications in medicine, and fundamental machine learning concepts. For instance, a participant described acquiring “a basic overview on AI, prompt generation, and machine learning.” Several participants also emphasized exposure to specific machine learning techniques, such as “oversampling, undersampling, and Shapley Additive Explanations plots.”

Development of Ethical and Reflective Awareness in AI Use

Beyond technical skills, participants demonstrated an increased awareness of ethical considerations and the reflective use of AI tools. Some participants mentioned learning about “AI ethics in medicine.” In addition, they reflected on how AI could support ongoing learning, with a participant noting learning “how to use AI tools to help with producing workable code and what concepts I need to be reading about for future self-learning.”

Learning Experience and Suggestions

Overall, students expressed positive perceptions of the learning environment, particularly valuing the instructors’ support and interactive teaching approach. Participants also offered constructive suggestions for improvement, such as covering “practical aspects of AI that can be applied to research projects.”

All instructors indicated that the “AI in Medicine” training workshop was well received, engaged students, and largely achieved its learning objectives, although some complex content required additional time. While participants enjoyed the workshop, their confidence in applying AI remained limited. For example, an instructor noted:

Participants demonstrated strong engagement and interest throughout the workshops. They consistently asked questions, ... , they found some of the machine learning content challenging. They asked numerous questions related to the depth of the ML material, which suggests that time constraints limited their ability to fully absorb these concepts. Several participants expressed a desire for additional time to gain a deeper understanding of the machine learning and explainable AI methods covered in the sessions.

Instructors also offered several suggestions for future implementations, including providing hands-on and practical sessions (emphasizing experiential learning and group activities, as these are the most engaging and effective), simplifying and allocating more time (simplifying technical content for nontechnical participants and providing more time and clearer instructions for practical exercises), and changing content depth and focus (narrowing the range of machine learning methods and exploring key topics in greater depth).

These self-reported findings align with several topics identified in study 1’s framework, particularly the basic understanding of AI concepts and their application in health care.

Discussion

Principal Results

This study addresses a critical gap in global medical education: the absence of standardized medical AI curriculum frameworks [30,31]. By integrating the UNESCO AI competency framework with the Miller pyramid, we translated high-level AI literacy principles into clinically actionable competencies across 6 dimensions. The 6D framework provides a structured reference for competency development, spanning foundational AI understanding to ethical, patient-centered, and practice-integrated application.

Building on this framework, we translated these competencies into a structured training program and piloted a workshop. The pilot implementation of module 3 provided initial evidence of the program’s feasibility, although the self-reported learning experience and confidence required further evaluation.

Comparison With Prior Work

Extending UNESCO’s AI competency framework to the medical domain [17], we incorporated domain-specific needs, such as patient-centered decision-making, ethical considerations in clinical AI, and health equity. Consistent with prior studies, our findings reinforce the importance of integrating both technical (eg, AI applications) and nontechnical competencies (eg, ethics) [11,12,15,35,36]. Importantly, this study responds to the growing impact of GenAI in health care by integrating competencies for its responsible use. Although recent literature has discussed both the applications and risks of GenAI in clinical contexts [21-23], few studies have articulated the competencies required for medical professionals to use these technologies safely and effectively. The inclusion of GenAI competencies provides a forward-looking perspective that is currently insufficiently examined in existing research.

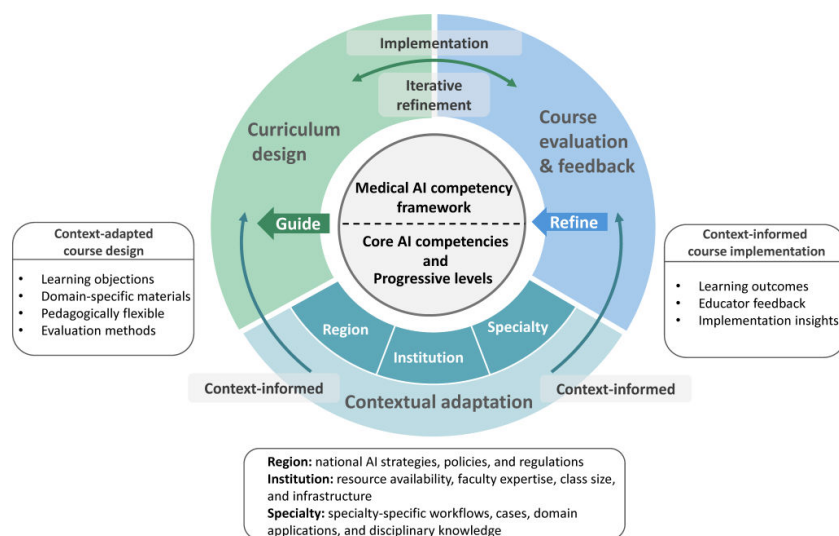
Prior studies on AI education in medicine have been largely domain-specific, focusing on areas such as machine learning fundamentals or programming [26,27,29,42]. For example, some programs emphasized conceptual understanding of AI methods through lectures [26,29], while others focused on developing basic programming skills [25]. Although these initiatives contribute to increased awareness and technical competence, they are typically organized around discrete topics rather than being guided by a unified framework. Furthermore, the literature has highlighted the absence of unified competency frameworks to guide these efforts [30,31]. In contrast, our study presents a framework-guided curriculum design approach that translates competency domains into training programs.

Theoretical and Educational Implications

This study provides both theoretical and practical implications. Theoretically, it proposes a medical AI competency framework that defines shared foundational AI competencies for medical education. It focuses on how medical professionals interact with, evaluate, communicate, and apply AI in health care, rather than replacing discipline-specific knowledge. We, therefore, view it as a living framework

(Figure 4). While it defines core competencies, its implementation should be tailored to regional, institutional, and specialty-specific contexts. Such adaptation occurs mainly at the curriculum level. For example, different specialties may require different cases, data, and learning materials. Feedback from course implementation can then support the iterative refinement of both the curriculum and the framework.

Figure 4. The medical artificial intelligence (AI) competency framework as a living framework.



Practically, the framework serves as a reference for curriculum design. By linking competency domains with the Miller levels, this framework supports curriculum mapping, the design of learning objectives and assessments, and the identification of training gaps. Additionally, prior work suggests that AI governance cannot rely solely on external constraints [9]. In this context, our framework supports the development of competencies for responsible AI use and may contribute to a broader governance-oriented approach to AI adoption in health care. This study further presents a competency-based AI training program that may serve as a model for other institutions. In addition, the explicit integration of GenAI into the competency framework and curriculum is particularly timely, given its rapid adoption in health care, education, and research.

Limitations and Future Directions

There are several limitations to this study. First, although the experts in this study represented diverse backgrounds, the majority were based in Singapore. As cultural attitudes toward AI adoption differ significantly across regions [63], the framework may mainly reflect Singapore's specific health care and educational context. The framework should be viewed as a living framework, open to ongoing refinement. Future research can validate and extend it using more diverse expert samples and combined deductive-inductive approaches. Future work may also extend the framework to specialty-specific applications (eg, course design and implementation) in areas such as radiology, surgery, and public health, using locally relevant cases, datasets, and assessment methods.

Second, the pilot was conducted as a voluntary workshop in Singapore, which may limit generalizability. Participants may have had greater prior interest in AI, introducing self-selection bias and potentially inflating the reported results. In addition, the pilot evaluation was based only on self-reported measures, covered only 1 module, and had a modest response rate (15/28), which may introduce nonresponse bias. Accordingly, these findings should be interpreted as preliminary evidence of feasibility. Future studies should evaluate the full program in broader settings using preassessments and postassessments of competency gains.

Third, while our program emphasizes real-world applicability, its short-term nature limits the direct assessment of AI competencies in clinical practice. Future work should collaborate with health care institutions to integrate training into clinical workflows and examine the longer-term impact across multiple levels, for example, using the Kirkpatrick model [64].

Conclusions

We developed a theoretically grounded medical AI competency framework. By translating competencies into a structured training program, it provides a practical application for integrating AI into medical education. Ultimately, this work aims to contribute to the development of an AI-ready health care workforce capable of using AI safely, ethically, and effectively in clinical practice.

Acknowledgments

The authors used ChatGPT for language polishing during manuscript preparation. All content was reviewed and verified by the authors. Responsibility for the final manuscript rests entirely with the authors. Generative artificial intelligence tools are not listed as authors and bear no responsibility for the final content.

Funding

This research is supported by the Ministry of Education, Singapore (grant: RG22/23, RS15/23, and LKCMedicine Start Up Grant).

Authors' Contributions

Conceptualization: C Cai

Data curation: C Cai, C Choi, LRW, XF

Formal analysis: C Cai, XF

Funding acquisition: XF

Investigation: C Cai

Methodology: C Cai

Resources: C Choi, LRW, GZ, SMZ, ON, JD, WKH, DMC, BTKL, FL, SL, VS

Supervision: XF

Visualization: C Cai

Writing—original draft: C Cai

Writing—review and editing: C Cai, XF, GZ, SMZ, ON, JD, WKH, DMC, BTKL, FL, SL, VS

Conflicts of Interest

None declared.

Multimedia Appendix 1

Initial medical artificial intelligence (AI) competency framework and codebook for deductive content analysis.

[\[DOCX File \(Microsoft Word File\), 27 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Course evaluation surveys and post-workshop surveys.

[\[DOCX File \(Microsoft Word File\), 15 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Deductive coding and Delphi results.

[\[DOCX File \(Microsoft Word File\), 23 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Final artificial intelligence (AI) training program.

[\[DOCX File \(Microsoft Word File\), 14 KB-Multimedia Appendix 4\]](#)

References

1. Singh P. Transforming healthcare through AI: enhancing patient outcomes and bridging accessibility gaps. SSRN. Preprint posted online on Jul 17, 2024. [doi: [10.2139/ssrn.5115767](https://doi.org/10.2139/ssrn.5115767)]
2. Kumar Y, Koul A, Singla R, Ijaz MF. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *J Ambient Intell Humaniz Comput*. 2023;14(7):8459-8486. [doi: [10.1007/s12652-021-03612-z](https://doi.org/10.1007/s12652-021-03612-z)] [Medline: [35039756](https://pubmed.ncbi.nlm.nih.gov/35039756/)]
3. Johnson KB, Wei WQ, Weeraratne D, et al. Precision medicine, AI, and the future of personalized health care. *Clin Transl Sci*. Jan 2021;14(1):86-93. [doi: [10.1111/cts.12884](https://doi.org/10.1111/cts.12884)] [Medline: [32961010](https://pubmed.ncbi.nlm.nih.gov/32961010/)]
4. Schork NJ. Artificial intelligence and personalized medicine. *Cancer Treat Res*. 2019;178:265-283. [doi: [10.1007/978-3-030-16391-4_11](https://doi.org/10.1007/978-3-030-16391-4_11)] [Medline: [31209850](https://pubmed.ncbi.nlm.nih.gov/31209850/)]
5. Pillai AS. AI-enabled hospital management systems for modern healthcare: an analysis of system components and interdependencies. *J Adv Anal Healthc Manag*. 2023;7(1):212-228. URL: <https://www.scribd.com/document/862060334/AI-enabled-Hospital-Management-Systems-for-Modern-Healthcare-An-Analysis-of-System-Components-and-Interdependencies> [Accessed 2026-05-06]
6. Alanazi A. Clinicians' views on using artificial intelligence in healthcare: opportunities, challenges, and beyond. *Cureus*. 2023;15(9):e45255. [doi: [10.7759/cureus.45255](https://doi.org/10.7759/cureus.45255)]

7. Esmailzadeh P. Challenges and strategies for wide-scale artificial intelligence (AI) deployment in healthcare practices: a perspective for healthcare organizations. *Artif Intell Med*. May 2024;151:102861. [doi: [10.1016/j.artmed.2024.102861](https://doi.org/10.1016/j.artmed.2024.102861)] [Medline: [38555850](https://pubmed.ncbi.nlm.nih.gov/38555850/)]
8. Meskó B, Hetényi G, Gyórfy Z. Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Serv Res*. Jul 13, 2018;18(1):545. [doi: [10.1186/s12913-018-3359-4](https://doi.org/10.1186/s12913-018-3359-4)] [Medline: [30001717](https://pubmed.ncbi.nlm.nih.gov/30001717/)]
9. Chen L. Beyond external constraints: the missing dimension of AI governance. SSRN. Preprint posted online on Mar 21, 2026. [doi: [10.2139/ssrn.6449738](https://doi.org/10.2139/ssrn.6449738)]
10. Aslam TM, Hoyle DC. Translating the machine: skills that human clinicians must develop in the era of artificial intelligence. *Ophthalmol Ther*. Feb 2022;11(1):69-80. [doi: [10.1007/s40123-021-00430-6](https://doi.org/10.1007/s40123-021-00430-6)] [Medline: [34807411](https://pubmed.ncbi.nlm.nih.gov/34807411/)]
11. Charow R, Jeyakumar T, Younus S, et al. Artificial intelligence education programs for health care professionals: scoping review. *JMIR Med Educ*. Dec 13, 2021;7(4):e31043. [doi: [10.2196/31043](https://doi.org/10.2196/31043)] [Medline: [34898458](https://pubmed.ncbi.nlm.nih.gov/34898458/)]
12. Lee J, Wu AS, Li D, Kulasegaram KM. Artificial intelligence in undergraduate medical education: a scoping review. *Acad Med*. Nov 1, 2021;96(11S):S62-S70. [doi: [10.1097/ACM.0000000000004291](https://doi.org/10.1097/ACM.0000000000004291)] [Medline: [34348374](https://pubmed.ncbi.nlm.nih.gov/34348374/)]
13. Liaw W, Kueper JK, Lin S, Bazemore A, Kakadiaris I. Competencies for the use of artificial intelligence in primary care. *Ann Fam Med*. 2022;20(6):559-563. [doi: [10.1370/afm.2887](https://doi.org/10.1370/afm.2887)] [Medline: [36443071](https://pubmed.ncbi.nlm.nih.gov/36443071/)]
14. Sand M, Durán JM, Jongsma KR. Responsibility beyond design: Physicians' requirements for ethical medical AI. *Bioethics*. Feb 2022;36(2):162-169. [doi: [10.1111/bioe.12887](https://doi.org/10.1111/bioe.12887)] [Medline: [34089625](https://pubmed.ncbi.nlm.nih.gov/34089625/)]
15. Singla R, Pupic N, Ghaffarizadeh SA, et al. Developing a Canadian artificial intelligence medical curriculum using a Delphi study. *NPJ Digit Med*. Nov 18, 2024;7(1):323. [doi: [10.1038/s41746-024-01307-1](https://doi.org/10.1038/s41746-024-01307-1)] [Medline: [39557985](https://pubmed.ncbi.nlm.nih.gov/39557985/)]
16. Weidener L, Fischer M. Artificial intelligence in medicine: cross-sectional study among medical students on application, education, and ethical aspects. *JMIR Med Educ*. Jan 5, 2024;10(1):e51247. [doi: [10.2196/51247](https://doi.org/10.2196/51247)] [Medline: [38180787](https://pubmed.ncbi.nlm.nih.gov/38180787/)]
17. AI competency framework for students. UNESCO. 2024. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000391105> [Accessed 2025-02-01]
18. Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. *Heliyon*. Feb 29, 2024;10(4):e26297. [doi: [10.1016/j.heliyon.2024.e26297](https://doi.org/10.1016/j.heliyon.2024.e26297)] [Medline: [38384518](https://pubmed.ncbi.nlm.nih.gov/38384518/)]
19. Nasir S, Khan RA, Bai S. Ethical framework for harnessing the power of AI in healthcare and beyond. *IEEE Access*. 2024;12:31014-31035. [doi: [10.1109/ACCESS.2024.3369912](https://doi.org/10.1109/ACCESS.2024.3369912)]
20. Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. *J Am Med Inform Assoc*. Mar 1, 2020;27(3):491-497. [doi: [10.1093/jamia/ocz192](https://doi.org/10.1093/jamia/ocz192)] [Medline: [31682262](https://pubmed.ncbi.nlm.nih.gov/31682262/)]
21. Chen A, Liu L, Zhu T. Advancing the democratization of generative artificial intelligence in healthcare: a narrative review. *J Hosp Manag Health Policy*. Jan 1, 2024;8:12-12. [doi: [10.21037/jhmhp-24-54](https://doi.org/10.21037/jhmhp-24-54)]
22. Moulaei K, Yadegari A, Baharestani M, Farzanbakhsh S, Sabet B, Reza Afrash M. Generative artificial intelligence in healthcare: a scoping review on benefits, challenges and applications. *Int J Med Inform*. Aug 2024;188:105474. [doi: [10.1016/j.ijmedinf.2024.105474](https://doi.org/10.1016/j.ijmedinf.2024.105474)] [Medline: [38733640](https://pubmed.ncbi.nlm.nih.gov/38733640/)]
23. Zhang P, Kamel Boulos MN. Generative AI in medicine and healthcare: promises, opportunities and challenges. *Future Internet*. 2023;15(9):286. [doi: [10.3390/fi15090286](https://doi.org/10.3390/fi15090286)]
24. Annappureddy R, Fornaroli A, Gatica-Perez D. Generative AI literacy: twelve defining competencies. *Digit Gov Res Pract*. Mar 31, 2025;6(1):1-21. [doi: [10.1145/3685680](https://doi.org/10.1145/3685680)]
25. Hu R, Fan KY, Pandey P, et al. Insights from teaching artificial intelligence to medical students in Canada. *Commun Med (Lond)*. 2022;2(1):63. [doi: [10.1038/s43856-022-00125-4](https://doi.org/10.1038/s43856-022-00125-4)] [Medline: [35668847](https://pubmed.ncbi.nlm.nih.gov/35668847/)]
26. Lindqwister AL, Hassanpour S, Lewis PJ, Sin JM. AI-RADS: an artificial intelligence curriculum for residents. *Acad Radiol*. Dec 2021;28(12):1810-1816. [doi: [10.1016/j.acra.2020.09.017](https://doi.org/10.1016/j.acra.2020.09.017)] [Medline: [33071185](https://pubmed.ncbi.nlm.nih.gov/33071185/)]
27. AI for healthcare. NUS. 2024. URL: <https://medicine.nus.edu.sg/continuing-education/course-catalogue/ai-for-healthcare> [Accessed 2024-01-20]
28. Oftring ZS, Deutsch K, Tolks D, Jungmann F, Kuhn S. Novel blended learning on artificial intelligence for medical students: qualitative interview study. *JMIR Med Educ*. May 26, 2025;11:e65220. [doi: [10.2196/65220](https://doi.org/10.2196/65220)] [Medline: [40418795](https://pubmed.ncbi.nlm.nih.gov/40418795/)]
29. Fundamentals of machine learning for healthcare. Stanford. URL: <https://online.stanford.edu/courses/som-xche0010-fundamentals-machine-learning-healthcare> [Accessed 2024-01-20]
30. Rincón EHH, Jimenez D, Aguilar LAC, Flórez JMP, Tapia ÁER, Peñuela CLJ. Mapping the use of artificial intelligence in medical education: a scoping review. *BMC Med Educ*. Apr 12, 2025;25(1):526. [doi: [10.1186/s12909-025-07089-8](https://doi.org/10.1186/s12909-025-07089-8)] [Medline: [40221725](https://pubmed.ncbi.nlm.nih.gov/40221725/)]
31. Tolentino R, Baradaran A, Gore G, Pluye P, Abbasgholizadeh-Rahimi S. Curriculum frameworks and educational programs in AI for medical students, residents, and practicing physicians: scoping review. *JMIR Med Educ*. Jul 18, 2024;10(1):e54793. [doi: [10.2196/54793](https://doi.org/10.2196/54793)] [Medline: [39023999](https://pubmed.ncbi.nlm.nih.gov/39023999/)]

32. Yao J. Research on the application risk and path of artificial intelligence in general education. *Adv Educ.* 2025;15(8):44-51. [doi: [10.12677/ae.2025.1581400](https://doi.org/10.12677/ae.2025.1581400)]
33. Cai C, Duell J, Chen DM, et al. Advancing AI literacy in medical education: a medical AI competency framework development. Presented at: Proceedings of the International Conference on Artificial Intelligence in Education; Jul 22-26, 2025; Palermo, Italy. [doi: [10.1007/978-3-031-98462-4_15](https://doi.org/10.1007/978-3-031-98462-4_15)]
34. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* Sep 1990;65(9 Suppl):S63-7. [doi: [10.1097/00001888-199009000-00045](https://doi.org/10.1097/00001888-199009000-00045)] [Medline: [2400509](https://pubmed.ncbi.nlm.nih.gov/2400509/)]
35. Paranjape K, Schinkel M, Nannan Panday R, Car J, Nanayakkara P. Introducing artificial intelligence training in medical education. *JMIR Med Educ.* Dec 3, 2019;5(2):e16048. [doi: [10.2196/16048](https://doi.org/10.2196/16048)] [Medline: [31793895](https://pubmed.ncbi.nlm.nih.gov/31793895/)]
36. Pupic N, Ghaffari-Zadeh A, Hu R, et al. An evidence-based approach to artificial intelligence education for medical students: a systematic review. *PLOS Digit Health.* Nov 2023;2(11):e0000255. [doi: [10.1371/journal.pdig.0000255](https://doi.org/10.1371/journal.pdig.0000255)] [Medline: [38011214](https://pubmed.ncbi.nlm.nih.gov/38011214/)]
37. Abràmoff MD, Tarver ME, Loyo-Berrios N, et al. Considerations for addressing bias in artificial intelligence for health equity. *NPJ Digit Med.* Sep 12, 2023;6(1):170. [doi: [10.1038/s41746-023-00913-9](https://doi.org/10.1038/s41746-023-00913-9)] [Medline: [37700029](https://pubmed.ncbi.nlm.nih.gov/37700029/)]
38. Thomasian NM, Eickhoff C, Adashi EY. Advancing health equity with artificial intelligence. *J Public Health Policy.* Dec 2021;42(4):602-611. [doi: [10.1057/s41271-021-00319-5](https://doi.org/10.1057/s41271-021-00319-5)] [Medline: [34811466](https://pubmed.ncbi.nlm.nih.gov/34811466/)]
39. Sim JH, Abdul Aziz YF, Mansor A, Vijayanathan A, Foong CC, Vadivelu J. Students' performance in the different clinical skills assessed in OSCE: what does it reveal? *Med Educ Online.* 2015;20(1):26185. [doi: [10.3402/meo.v20.26185](https://doi.org/10.3402/meo.v20.26185)] [Medline: [25697602](https://pubmed.ncbi.nlm.nih.gov/25697602/)]
40. Witheridge A, Ferns G, Scott-Smith W. Revisiting Miller's pyramid in medical education: the gap between traditional assessment and diagnostic reasoning. *Int J Med Educ.* Oct 25, 2019;10:191-192. [doi: [10.5116/ijme.5d9b.0c37](https://doi.org/10.5116/ijme.5d9b.0c37)] [Medline: [31655795](https://pubmed.ncbi.nlm.nih.gov/31655795/)]
41. Emam KE. Benchmarking Kappa: interrater agreement in software process assessments. *Empir Softw Eng.* Jun 1999;4(2):113-133. [doi: [10.1023/A:1009820201126](https://doi.org/10.1023/A:1009820201126)]
42. Abid A, Murugan A, Banerjee I, Purkayastha S, Trivedi H, Gichoya J. AI education for fourth-year medical students: two-year experience of a web-based, self-guided curriculum and mixed methods study. *JMIR Med Educ.* Feb 20, 2024;10:e46500. [doi: [10.2196/46500](https://doi.org/10.2196/46500)] [Medline: [38376896](https://pubmed.ncbi.nlm.nih.gov/38376896/)]
43. Huang CJ, Wu T, Lu JT, et al. Developing a medical artificial intelligence course for high school students. Presented at: International Forum on Medical Imaging in Asia 2021; Jan 24-27, 2021:103-108; Taipei, Taiwan. [doi: [10.1117/12.2590769](https://doi.org/10.1117/12.2590769)]
44. Laupichler MC, Hadizadeh DR, Wintergerst MWM, et al. Effect of a flipped classroom course to foster medical students' AI literacy with a focus on medical imaging: a single group pre-and post-test study. *BMC Med Educ.* Nov 18, 2022;22(1):803. [doi: [10.1186/s12909-022-03866-x](https://doi.org/10.1186/s12909-022-03866-x)] [Medline: [36397110](https://pubmed.ncbi.nlm.nih.gov/36397110/)]
45. Yao J, Li M, Yang B. Blended learning into database courses: a path to integrating online and offline resources. Presented at: 2025 16th International Conference on E-Education, E-Business, E-Management and E-Learning (IC4e); Jan 17-19, 2025:377-381; Tokyo, Japan. [doi: [10.1109/IC4e65071.2025.11075421](https://doi.org/10.1109/IC4e65071.2025.11075421)]
46. Kolb DA. *Experiential Learning: Experience as the Source of Learning and Development.* Prentice Hall; 1984. ISBN: 9780132952613
47. Gibbs G. *Learning by Doing: A Guide to Teaching and Learning Methods.* Further Education Unit; 1988. ISBN: 9781853380716
48. Turoff M. *The Delphi Method: Techniques and Applications.* Addison-Wesley; 1975. ISBN: 9780201042948
49. Shariff NJ. Utilizing the Delphi survey approach: a review. *J Nurs Care.* 2015;04(3):2167-1168. [doi: [10.4172/2167-1168.1000246](https://doi.org/10.4172/2167-1168.1000246)]
50. Reeves G, Jauch LR. Curriculum development through Delphi. *Res High Educ.* 1978;8(2):157-168. [doi: [10.1007/BF00992116](https://doi.org/10.1007/BF00992116)]
51. Sitlington H, Coetzer A. Using the Delphi technique to support curriculum development. *Educ + Train.* Apr 13, 2015;57(3):306-321. [doi: [10.1108/ET-02-2014-0010](https://doi.org/10.1108/ET-02-2014-0010)]
52. York CS, Ertmer PA. Towards an understanding of instructional design heuristics: an exploratory Delphi study. *Educ Technol Res Dev.* Dec 2011;59(6):841-863. [doi: [10.1007/s11423-011-9209-2](https://doi.org/10.1007/s11423-011-9209-2)]
53. Taylor E. We agree, don't we? The Delphi method for health environments research. *HERD.* Jan 2020;13(1):11-23. [doi: [10.1177/1937586719887709](https://doi.org/10.1177/1937586719887709)] [Medline: [31887097](https://pubmed.ncbi.nlm.nih.gov/31887097/)]
54. Zheng J, Lou L, Xie Y, et al. Model construction of medical endoscope service evaluation system-based on the analysis of Delphi method. *BMC Health Serv Res.* Jul 9, 2020;20(1):629. [doi: [10.1186/s12913-020-05486-x](https://doi.org/10.1186/s12913-020-05486-x)] [Medline: [32646429](https://pubmed.ncbi.nlm.nih.gov/32646429/)]

55. Banayan J, Blood A, Park YS, Shahul S, Scavone BM. A modified Delphi method to create a scoring system for assessing team performance during maternal cardiopulmonary arrest. *Hypertens Pregnancy*. 2015;34(3):314-331. [doi: [10.3109/10641955.2015.1033926](https://doi.org/10.3109/10641955.2015.1033926)] [Medline: [25954824](https://pubmed.ncbi.nlm.nih.gov/25954824/)]
56. Eubank BH, Mohtadi NG, Lafave MR, et al. Using the modified Delphi method to establish clinical consensus for the diagnosis and treatment of patients with rotator cuff pathology. *BMC Med Res Methodol*. May 20, 2016;16:56. [doi: [10.1186/s12874-016-0165-8](https://doi.org/10.1186/s12874-016-0165-8)] [Medline: [27206853](https://pubmed.ncbi.nlm.nih.gov/27206853/)]
57. Jacob E, Duffield C, Jacob D. A protocol for the development of a critical thinking assessment tool for nurses using a Delphi technique. *J Adv Nurs*. Aug 2017;73(8):1982-1988. [doi: [10.1111/jan.13306](https://doi.org/10.1111/jan.13306)] [Medline: [28329414](https://pubmed.ncbi.nlm.nih.gov/28329414/)]
58. Hou YF, Liu D, Zhou SG, et al. The capacity for psychotherapy process scale: development and psychometric validation of patient self-reports. *Psychother Res*. Feb 2021;31(2):236-246. [doi: [10.1080/10503307.2020.1769218](https://doi.org/10.1080/10503307.2020.1769218)] [Medline: [32496976](https://pubmed.ncbi.nlm.nih.gov/32496976/)]
59. Franklin AE, Burns P, Lee CS. Psychometric testing on the NLN student satisfaction and self-confidence in learning, simulation design scale, and educational practices questionnaire using a sample of pre-licensure novice nurses. *Nurse Educ Today*. Oct 2014;34(10):1298-1304. [doi: [10.1016/j.nedt.2014.06.011](https://doi.org/10.1016/j.nedt.2014.06.011)] [Medline: [25066650](https://pubmed.ncbi.nlm.nih.gov/25066650/)]
60. Reeve J, Tseng CM. Agency as a fourth aspect of students' engagement during learning activities. *Contemp Educ Psychol*. Oct 2011;36(4):257-267. [doi: [10.1016/j.cedpsych.2011.05.002](https://doi.org/10.1016/j.cedpsych.2011.05.002)]
61. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. Jan 2006;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
62. Hu Y. Toward retrieval-grounded evaluation for conversational large language model-based risk assessment. *JMIR AI*. Mar 12, 2026;5:e90759. [doi: [10.2196/90759](https://doi.org/10.2196/90759)] [Medline: [41818631](https://pubmed.ncbi.nlm.nih.gov/41818631/)]
63. Yam KC, Tan T, Jackson JC, Shariff A, Gray K. Cultural differences in people's reactions and applications of robots, algorithms, and artificial intelligence. *Manag Organ Rev*. Oct 2023;19(5):859-875. [doi: [10.1017/mor.2023.21](https://doi.org/10.1017/mor.2023.21)]
64. Introduction to the New World Kirkpatrick® model. Kirkpatrick Partners; 2015. URL: <https://www.kirkpatrickpartners.com/wp-content/uploads/2021/11/Introduction-to-The-New-World-Kirkpatrick-Model-1.pdf> [Accessed 2026-05-06]

Abbreviations

AI: artificial intelligence

AS: agreement score

FS: full score

GenAI: generative artificial intelligence

HIPAA: Health Insurance Portability and Accountability Act

UNESCO: United Nations Educational, Scientific and Cultural Organization

Edited by Alicia Stone; peer-reviewed by Mohammed Rohaim, Monzurul Roni, Yihan Hu; submitted 09.Jan.2026; final revised version received 20.Apr.2026; accepted 22.Apr.2026; published 20.May.2026

Please cite as:

Cai C, Zhu G, Zhou SM, Ng O, Duell J, Ho WK, Chen DM, Lee B, Li F, Liu S, Sudarshan V, Wang LR, Choi C, Fan X

A Competency Framework for Medical AI Education: Mixed Methods Study

JMIR Med Educ 2026;12:e91116

URL: <https://mededu.jmir.org/2026/1/e91116>

doi: [10.2196/91116](https://doi.org/10.2196/91116)

© Chang Cai, Gaoxia Zhu, Shang-Ming Zhou, Olivia Ng, Jamie Duell, Weng Kin Ho, Daisy Minghui Chen, Bennett Lee, Fang Li, Siyuan Liu, Vidya Sudarshan, Li Rong Wang, Chanwoo Choi, Xiuyi Fan. Originally published in *JMIR Medical Education* (<https://mededu.jmir.org>), 20.May.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Medical Education*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mededu.jmir.org/>, as well as this copyright and license information must be included.