

Expert Consensus on Artificial Intelligence Proficiency for Medical Educators (2025 Edition)

Mengchun Gong

Chinese Academy of Medical Sciences & Peking Union Medical College

Hui Pan

Chinese Academy of Medical Sciences & Peking Union Medical College

Jinghui Lu

Guangdong Medical University

Zhirong Zeng

Guangdong Medical University

Wei Chen

Sun Yat-sen University

Hao Wu

The First Affiliated Hospital of Chongqing Medical University

Li Luo

Fudan University

Wenwen Sun

Chinese Academy of Medical Sciences & Peking Union Medical College

Hui Liu

Chinese Academy of Medical Sciences & Peking Union Medical College

Hang Li

Peking Union Medical College Hospital

Wei Wang

Chinese Academy of Medical Sciences & Peking Union Medical College

Yan Luo

Chinese Academy of Medical Sciences & Peking Union Medical College

Bohan Zhang

Guangdong Medical University

Xunming Ji

jixunming@vip.163.com

Chinese Academy of Medical Sciences & Peking Union Medical College

Keywords: Medical education, Faculty development, Artificial intelligence literacy, Competency framework, Expert consensus

Posted Date: October 7th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10925679/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

In response to the challenges that the profound integration of generative artificial intelligence (AI) into medical education poses for educators' roles and competencies, this consensus used a systematic literature review, preliminary framework development, multiple rounds of expert preparatory work, and a structured Delphi method to extensively solicit the views of 60 interdisciplinary experts and to examine and calibrate the core competency items and observable standards of medical educators' AI proficiency. It proposes a clear, medically distinctive, forward-looking, and operational framework of AI proficiency for medical educators. The framework comprises five dimensions and 25 competency items: Value Cognition and Ethical Foundation; Technical Understanding and Tool Application; Teaching Integration and Innovative Practice; Learning Assessment and Precision Empowerment; and Professional Development and Ecosystem Co-construction. The items are classified into 11 foundational competencies required of all medical educators and 14 developmental competencies for those pursuing excellence. Each item is described through its scope and definition and key behavioral indicators, with corresponding observable assessment indicators. The consensus provides a scientific basis for the professional development of medical educators and institutional faculty capacity building and establishes a key reference for educator competency development during the digital transformation of medical education. This article is an authorized English secondary publication of the consensus originally published in Chinese.

Introduction

Medical education is undergoing a historic transformation driven by emerging technologies such as generative AI. In this context, the role of educators must expand beyond mere knowledge transmitters to become integrators of AI technologies, designers of pedagogical innovation, guardians of ethical standards, and co-creators of educational ecosystems¹⁻². However, faculty currently face three major challenges: (1) uneven development of AI competencies, with usage often limited to superficial tool application; (2) insufficient ethical awareness, particularly regarding systemic understanding of data privacy and algorithmic fairness; (3) the absence of standardized evaluation criteria to scientifically assess the multifaceted competencies required in the AI era³.

Against this backdrop, China's Ministry of Education issued the *Guidelines for the Application of Generative Artificial Intelligence in Teaching (Version 1)* in November 2025⁴. These guidelines outline six key application domains and 30 specific scenarios, aiming to shift educators' focus from passive technology adoption toward human-AI collaborative teaching. They reinforce the teacher's central role in intelligent education and embed ethical principles into practice, offering crucial guidance for educational implementation. Nevertheless, medical education—characterized by its high degree of specialization, hands-on clinical orientation, and acute ethical sensitivity—presents unique complexities, including intricate teaching scenarios involving patient data and clinical decision-making. These features impose more specialized and systematic demands on faculty AI literacy.

Therefore, building upon the national guidelines, there is an urgent need to develop a professional AI literacy framework tailored specifically to medical education—one that is grounded in China’s national context, aligned with international standards, and closely attuned to the realities of medical teaching. This consensus statement responds directly to that need by addressing the core question: “What AI literacy should future medical educators possess?” It aims to provide a scientific reference for faculty development, institutional management, and policy formulation. As a proactive response to China’s national strategy for educational digitalization, this consensus supports the broader goals of building a strong education system and advancing the “New Medical Education” initiative. Drawing on cutting-edge international concepts, it seeks to cultivate future physicians who integrate AI capabilities with humanistic values, thereby contributing Chinese insights and enhancing the global influence of China’s medical education.

This consensus primarily targets faculty engaged in teaching across disciplines—including clinical medicine, basic medical sciences, public health, nursing, dentistry, and related fields—at medical schools, affiliated teaching hospitals, and other clinical training institutions nationwide (including physician-educators who concurrently provide clinical care). Its scope encompasses medical universities, affiliated hospitals, other medical teaching units, faculty development centers, and medical education administration and evaluation bodies. The consensus serves multiple purposes: (1) as a self-assessment and professional development reference for individual educators; (2) as a scientific basis for institutions to design AI-related faculty training, build teaching platforms, conduct performance evaluations, manage promotions, and recognize teaching excellence; and (3) as expert guidance for education authorities in formulating relevant standards and policies. Ultimately, it aims to foster a medical AI education ecosystem characterized by “value-driven leadership, technology-enabled empowerment, pedagogical innovation, evidence-based evaluation, and collaborative ecology.”

In the context of intelligent transformation in medical education, AI literacy for medical educators is defined as an integrated professional competency—a composite of knowledge, skills, attitudes, and values encompassing cognitive understanding, technical application, ethical judgment, innovative practice, and ecosystem-building capacities necessary for thriving in a human-AI collaborative paradigm. Its ultimate goal is to empower educators to become “wise educators” who can understand, harness, and ethically leverage AI to enhance teaching and learning ⁵⁻⁶.

To translate this overarching literacy into concrete, observable, and assessable developmental objectives, this consensus adopts a competency-based approach to construct the Competency Items of Medical Educators’ AI Proficiency (CAIP-ME) framework. CAIP-ME comprises a set of well-defined, non-overlapping, and clearly bounded competency items. Each item is articulated through a precise “definition of scope” and associated “key behavioral indicators,” accompanied by suggested assessment approaches. The framework further distinguishes between foundational competencies—essential for all educators to ensure baseline teaching quality, safety, and ethical integrity—and advanced developmental competencies that reflect personalized growth trajectories toward teaching excellence and innovation leadership. Thus, the CAIP-ME framework operationalizes the abstract concept of AI literacy into a

structured, actionable roadmap, providing a clear pathway for systematic faculty development and precise evaluation.

1. Consensus Development Methods and Process

1.1 Literature Search and Evidence Integration

A systematic search was conducted across databases including PubMed, Web of Science, CNKI (China National Knowledge Infrastructure), and Wanfang, focusing on the period from 2020 to 2025. The search targeted topics such as "medical educators" and "AI literacy", resulting in the inclusion of 47 Chinese and English articles. Core competency dimensions, assessment criteria, and practical experiences were distilled from these sources. Based on the teacher competency model (knowledge-skills-attitudes), AI literacy was dissected into observable and quantifiable competencies. Following this, 24 cross-disciplinary experts participated in online seminars to preliminarily validate and refine these competencies, culminating in the CAIP-ME framework encompassing five dimensions and 25 specific competencies, laying the groundwork for subsequent Delphi consensus development. Supporting search and screening materials are provided in Additional files 1–3.

1.2 Consensus Building Using the Delphi Method

An enhanced Delphi method was employed through two rounds of structured expert consultations to build consensus. Sixty experts from universities, research institutes, and teaching hospitals formed a multidisciplinary panel. Predominantly senior-level with high academic qualifications (91% PhDs, 86% full professors), the panel covered areas like clinical medical education (68%), educational technology/AI (40%), and educational management/evaluation (28%), achieving an overall authority coefficient of 0.849. In the first round, all 25 competencies were rated on a 5-point scale for importance, scoring between 4.105 and 4.877, ensuring their retention. The second round focused on categorizing these competencies as either foundational or developmental. Ultimately, 11 foundational and 14 developmental competencies were identified, reflecting a hierarchical development philosophy centered on ethics-first, application-driven, and ecosystem co-construction principles. Supporting panel, consultation, meeting, and item-level Delphi materials are provided in Additional files 4–10; the completed web-survey reporting checklist is provided in Additional file 13.

1.3 Framework Design of CAIP-ME

The construction of this consensus framework is grounded in teacher competency models, constructivist pedagogy, and evidence-based medical education theories, closely aligning with the theoretical-practical-ethical triad characteristic of medical education. It ensures that competency dimensions and assessment methods adhere to educational principles while addressing the specificity, practicality, and ethical sensitivity of medical instruction. Emphasizing practical applicability, all competencies are derived from real-world medical teaching scenarios and include clear quantitative evaluation criteria for objective and feasible assessment. Additionally, innovatively distinguishing

between "Foundational (F)" and "Developmental (D)" levels accommodates diverse institutional, disciplinary, and career-stage needs, maintaining standardization while allowing flexibility. The complete competency table and evaluation suggestions are provided in Additional file 11.

Core concepts guiding the framework design include: (1) Prioritizing ethics and humanistic values in technology use; (2) Leveraging technology to enhance rather than replace human intelligence and emotional connections; (3) Promoting interdisciplinary integration and systematic advancement through a value-technology-teaching-evaluation-development pentagonal structure; and (4) Adapting international advanced concepts to address China's medical education challenges, aiming to establish a globally communicative "Chinese Solution".

Based on these principles, the CAIP-ME framework consists of five interconnected, progressively layered core dimensions: The development pathway, final structure, and intended use of the CAIP-ME framework are summarized in Fig. 1.

1. *Value Cognition and Ethical Foundation (CAIP-ME 1–4)*: Establishes value orientation and ethical guidelines for technology use, addressing fundamental questions of why and how to ethically apply AI ⁷.
2. *Technical Understanding and Tool Application (CAIP-ME 5–9)*: Provides the foundational skills needed to understand and safely utilize AI tools.
3. *Teaching Integration and Innovative Practice (CAIP-ME 10–14)*: Focuses on creatively integrating AI into curriculum design to foster student-centered intelligent learning environments.
4. *Learning Assessment and Precision Empowerment (CAIP-ME 15–18)*: Utilizes AI for multi-dimensional dynamic assessment and personalized support to promote student development.
5. *Professional Development and Ecosystem Co-construction (CAIP-ME 19–25)*: Encourages continuous professional growth and active participation in building supportive AI education innovation cultures.

2. Detailed Elaboration of the 25 Competencies Under the Five Dimensions (CAIP-ME)

2.1 Value Cognition and Ethical Foundation (Dimension One)

2.1.1 Recognition of the Value of AI-Enabled Medical Education (CAIP-ME1)

- **Scope and Definition:** Profoundly understands the strategic value of AI as a transformative tool for improving the quality of medical education, promoting educational equity, and preparing future physicians. Moves beyond a tool-centered perspective to recognize, from the standpoint of

modernization in medical education and the Healthy China strategy, the necessity and urgency of developing one's own AI proficiency.

- **Key Behavioral Indicators:** In teaching seminars or course design, proactively explains the potential value of AI for addressing specific challenges in medical education, such as teaching rare diseases or training for high-risk procedures; consciously plans and invests time and effort in learning and improving AI-enabled teaching competencies; and actively advocates within the team for appropriate pathways to integrate AI with teaching.

2.1.2 Identification and Assessment of Ethical Risks in Medical AI (CAIP-ME2)

- **Scope and Definition:** Systematically identifies ethical, legal, and social risks that may arise when AI is used in medical education, research, and simulated clinical applications, including but not limited to breaches of data privacy, algorithmic bias and discrimination, ambiguity in the attribution of responsibility, alienation of the physician-patient relationship, and technological dependence. On this basis, understands the core requirements of AI governance, including transparency (explainability), accountability (clear attribution of responsibility), and auditability (traceability of processes), and makes preliminary judgments in accordance with the principles of medical ethics, including beneficence, non-maleficence, justice, and autonomy.
- **Key Behavioral Indicators:** Before using any AI tool involving patient data or students' personal information, proactively reviews its privacy policy, data-governance framework, and regulatory compliance; critically analyzes an AI medical case or teaching tool, identifies possible fairness concerns such as inadequate data representativeness, and questions the transparency and auditability of its decision logic; and, when designing teaching activities, fully considers recommendations in relevant guidance documents and anticipates potential ethical disputes and corresponding responsibilities ⁸.

2.1.3 Guidance on Academic Integrity and Standards for AI Use (CAIP-ME3)

- **Scope and Definition:** Clearly defines the boundaries and standards for the appropriate use of AI in medical learning, assignments, theses, and research. Establishes clear rules and effectively guides students to use AI-assisted learning appropriately and transparently, while resolutely preventing and investigating academic misconduct such as plagiarism, fabrication, and improper attribution.
- **Key Behavioral Indicators:** Develops and publishes explicit Rules for the Use of AI-Assisted Learning for the courses taught, specifying the circumstances in which AI may be used and how its use should be acknowledged; demonstrates preliminary awareness and ability to identify AI-generated content when assessing student assignments, including through oral questioning; and provides serious education and correction when students use AI improperly.

2.1.4 Ability to Promote Equity and Justice in Educational Practice (CAIP-ME4)

- **Scope and Definition:** Recognizes and responds proactively to inequities in AI-enabled education caused by the digital divide, unequal resources, and algorithmic bias. Takes active measures in teaching practice to ensure that all students, particularly those with weaker foundations or limited resources, can benefit equitably from AI-enabled education.
- **Key Behavioral Indicators:** Ensures that core AI teaching resources, including software and platform access, are available to all enrolled students free of charge and without barriers; provides additional guidance or alternative learning pathways for students who are not proficient in the relevant technologies; and consciously considers diversity in disease spectra and population characteristics when selecting or designing AI teaching cases.

2.2 Technical Understanding and Tool Application (Dimension Two)

2.2.1 Understanding of Core AI Concepts and Medical Applications (CAIP-ME5)

- **Scope and Definition:** Understands the basic principles and capability boundaries of core concepts such as machine learning, generative AI, and large language models. Clearly explains typical applications of AI in medical settings, including imaging diagnosis, medical-record analysis, drug discovery, and health management, as well as their limitations.
- **Key Behavioral Indicators:** Explains to students in accessible language why AI sometimes makes implausible errors, including hallucinations; compares the advantages and disadvantages of AI-assisted diagnosis and human expert diagnosis in different situations; and determines whether a particular AI technology genuinely matches the essential needs of the teaching objectives when selecting a teaching project.

2.2.2 Evaluation and Selection of Medical AI Tools (CAIP-ME6)

- **Scope and Definition:** Based on specific teaching objectives, course content, and disciplinary characteristics, retrieves information on, compares the functions of, reviews the compliance of, and evaluates the applicability of diverse medical AI tools, including teaching tools, research tools, and clinical decision-support tools, and makes a reasoned selection.
- **Key Behavioral Indicators:** For a course being taught, screens several AI-assisted clinical-diagnosis or educational-practice tools and selects a set suitable for teaching demonstrations, explaining the reasons for the selection; reads and understands provisions on data ownership and privacy in tool

user agreements; and understands the pricing of key tools, open-source alternatives, and any institutional group licenses that may be available.

2.2.3 Operational Competency in Data Security and Privacy Protection (CAIP-ME7)

- **Scope and Definition:** Establishes a strong awareness of data security and proficiently applies standardized processes and technical methods, such as de-identification, when handling sensitive data in teaching and research. Strictly complies with applicable laws and regulations, including the Personal Information Protection Law, to ensure data security throughout the full data lifecycle.
- **Key Behavioral Indicators:** Before publishing a teaching case or research paper, performs standardized de-identification of any patient information; appropriately safeguards electronic documents containing student or patient information and uses encrypted storage and transmission; and correctly uses secure institutional platforms for data storage and analysis.

2.2.4 AI-Assisted Literature Retrieval and Analysis (CAIP-ME8)

- **Scope and Definition:** Uses AI-driven academic search engines, reference-management tools, and evidence-synthesis platforms proficiently to conduct literature retrieval, batch screening, key-point synthesis, and trend analysis efficiently, thereby substantially improving the depth and breadth of literature research on a specific topic.
- **Key Behavioral Indicators:** Uses AI tools to rapidly map research hotspots and key publications from the previous three years in a medical field, such as applications of AI in ophthalmology; compares differences among the perspectives of multiple papers and generates a literature-comparison matrix; and effectively integrates the results of AI-assisted literature analysis into updates of teaching plans or the rationale for research projects.

2.2.5 Basic Processing and Visualization of Research Data (CAIP-ME9)

- **Scope and Definition:** Uses AI-enhanced data-analysis tools or programming environments to perform basic cleaning, transformation, statistical analysis, and graphical presentation of structured medical research data, thereby improving research efficiency.
- **Key Behavioral Indicators:** Uses tools to automatically detect and handle missing and abnormal values in a dataset; selects appropriate statistical graphics, such as box plots or heat maps, according to the data type and research question and produces visualizations; and interprets preliminary statistical reports generated by AI tools.

2.3 Teaching Integration and Innovative Practice (Dimension Three)

2.3.1 AI-Enhanced Instructional Design (CAIP-ME10)

- **Scope and Definition:** Uses technologies such as generative AI and knowledge graphs to support learner analysis, goal setting, activity design, and resource matching, and develops personalized and highly interactive blended-learning solutions to optimize instructional strategies.
- **Key Behavioral Indicators:** Uses AI tools and data from previous cohorts to predict potential learning difficulties in different chapters of the current course; designs AI virtual standardized-patient interview tasks as pre-class preparation; and creates differentiated AI-recommended reading lists or exercise banks for students with different levels of learning progress.

2.3.2 Application of Intelligent Virtual Simulation in Teaching (CAIP-ME11)

- **Scope and Definition:** Leads or participates substantially in the development and use of intelligent clinical-skills training environments based on virtual reality, augmented reality, or high-fidelity patient simulators. Uses these environments to teach high-risk, high-cost, or rare clinical scenarios and provides effective guidance and debriefing throughout the learning process.
- **Key Behavioral Indicators:** Configures complex emergencies, such as intraoperative bleeding, in a virtual-surgery training system and guides students in responding; uses AI-driven virtual standardized patients for longitudinal history-taking instruction and retrieves dialogue records for targeted critique; and develops complex scenarios, such as emergency responses to sudden public-health events, into virtual-simulation training modules.

2.3.3 AI-Assisted Creation of Teaching Resources (CAIP-ME12)

- **Scope and Definition:** Uses generative AI tools effectively to produce high-quality customized teaching materials, including anatomical illustrations, schematic images of pathological sections, animations of disease mechanisms, scripts for typical cases, and exercises of different levels of difficulty.
- **Key Behavioral Indicators:** Uses AI to generate a set of comparative images showing different stages of disease, such as diabetic retinopathy or cervical cancer; converts a complex procedural guide, such as cardiopulmonary resuscitation, into a clear step-by-step graphic or short-video script; and ensures that AI-generated teaching resources satisfy requirements for medical accuracy and indicate their source or mode of generation.

2.3.4 Organization of Human-AI Collaborative Classroom Teaching (CAIP-ME13)

- **Scope and Definition:** Integrates AI tools, such as real-time question-answering agents, collaborative whiteboards, and intelligent polling systems, appropriately into in-person or online classroom processes; dynamically coordinates teacher instruction, AI assistance, and student activity; and creates a highly participatory and efficient classroom environment.
- **Key Behavioral Indicators:** Introduces an AI conversational assistant during teaching to answer students' conceptual questions in real time while the teacher concentrates on deeper discussion; uses intelligent polling tools to assess students' immediate understanding of a diagnostic point and adjusts the pace of teaching accordingly; and manages the classroom effectively so that technology serves the learning objectives rather than distracting students.

2.3.5 Integration of Clinical AI Tools into Teaching Scenarios (CAIP-ME14)

- **Scope and Definition:** In authentic clinical work, such as ward rounds, outpatient teaching, or observation of surgical procedures, naturally transforms the use of clinical AI decision-support systems and intelligent prompts in electronic medical records into teaching opportunities, guiding students to understand the clinical reasoning involved in human-AI collaboration.
- **Key Behavioral Indicators:** During ward rounds, demonstrates how AI-assisted tools can be used to rapidly review current diagnostic and treatment guidelines and relates this information to the patient's actual condition; discusses with students the rationale and limitations of differential-diagnosis suggestions generated by an AI system; and uses real cases to explain how clinicians decide whether to override or accept an AI alert.

2.4 Learning Evaluation and Precision Empowerment (Dimension Four)

2.4.1 Design of AI-Supported Formative Assessment (CAIP-ME15)

- **Scope and Definition:** Uses intelligent question banks, online assessment platforms, and related technologies to design and implement formative assessment throughout the teaching process. Continuously collects learning-process data through embedded knowledge checks and chapter exercises.
- **Key Behavioral Indicators:** Configures associated instant exercises for each video-based knowledge point on an online course platform; designs sequential assessment tasks based on cases, with AI providing different subsequent scenarios according to the student's previous choice; and regularly,

such as weekly, releases unit quizzes assembled by AI and provides immediate feedback on the results.

2.4.2 Interpretation of Learning Data and Instructional Intervention (CAIP-ME16)

- **Scope and Definition:** Accurately interprets learning-analysis reports generated by AI-based assessment systems, such as overviews of class-level knowledge mastery and analyses of individual learning pathways; identifies generally weak teaching components and individualized learning difficulties; and implements effective group-level remediation or individualized learning recommendations accordingly.
- **Key Behavioral Indicators:** Uses system reports showing a high error rate for particular teaching content to arrange a focused review session; privately sends encouraging messages and links to key tutorial resources to students whose progress is substantially delayed; and distinguishes whether learning-data findings reflect problems in knowledge comprehension or problems in the wording of test items.

2.4.3 Diagnosis and Support of Students' Digital Literacy (CAIP-ME17)

- **Scope and Definition:** Observes and evaluates students' foundational ability to use multimodal data, machine-learning algorithms, and other digital tools for learning, including information literacy and technical operational skills. Identifies students experiencing difficulties with information technology or data science and provides the necessary skills guidance or alternative pathways to promote equity at the starting point.
- **Key Behavioral Indicators:** Conducts a simple questionnaire or survey at the beginning of a course on students' use of information-science and data tools; provides demonstrations or peer guidance for students unfamiliar with basic data-analysis software; and permits students who cannot complete a particular online task because of technical constraints to submit equivalent work offline.

2.4.4 Tracking and Evaluation of Students' AI Proficiency Development (CAIP-ME18)

- **Scope and Definition:** In accordance with the Expert Consensus on the Artificial Intelligence Proficiency Competency List and Assessment Framework for Medical Students (2025 Edition), comprehensively understands the specific connotations and assessment requirements of students' AI proficiency, designs observation points or assessment tasks within the curriculum, and promotes the steady development of their AI proficiency⁹.

- **Key Behavioral Indicators:** Based on that framework, designs authentic tasks embedded in the curriculum, such as clinical decision-making analyses based on AI-assisted diagnostic reports; provides targeted feedback on students' appropriate use of AI tools, critical appraisal of information, and ethical decision-making; and uses relevant platforms to monitor and track the development of students' AI proficiency and provide targeted instructional guidance.

2.5 Professional Development and Ecosystem Co-construction (Dimension Five)

2.5.1 Reflection on and Optimization of AI-Enabled Teaching Practice (CAIP-ME19)

- **Scope and Definition:** Develops a habit of evidence- and data-informed reflection on teaching. Systematically collects evidence of the effectiveness of AI-enabled teaching, including student feedback, performance data, and peer observations, as well as indications of problems; analyzes the causes; and develops and implements feasible improvement plans to establish a continuous cycle of professional practice improvement.
- **Key Behavioral Indicators:** After trying a new AI teaching tool, writes a concise practice journal recording strengths and limitations; explicitly adjusts an ineffective AI-assisted component from the previous teaching cycle in the design of the current cycle and consciously compares outcomes; and willingly shares lessons learned from unsuccessful attempts with colleagues.

2.5.2 Interdisciplinary Collaboration and Communication (CAIP-ME20)

- **Scope and Definition:** Communicates and collaborates effectively with professionals from non-medical backgrounds, including educational technologists, AI engineers, data scientists, and ethicists. Accurately translates the needs of medical education into requirements for technology development and jointly advances innovative teaching projects.
- **Key Behavioral Indicators:** Collaborates with engineers to prepare a functional requirements document for developing an AI teaching tool, such as a virtual-acupuncture training system; explains the concrete manifestation of a clinical teaching problem in clear language during interdisciplinary project meetings; and respects and seeks to understand the professional logic and constraints of collaborators.

2.5.3 Research on AI-Enabled Educational Reform and Translation of Outcomes (CAIP-ME21)

- **Scope and Definition:** Proactively conducts action research or educational experiments on the application of AI in medical education and systematically organizes the lessons learned, whether successful or unsuccessful, into teaching cases, papers, or reports that can be shared and replicated, thereby contributing to the public knowledge base.
- **Key Behavioral Indicators:** Leads a university-level educational-reform project, such as exploring an AI-plus-VR teaching model for general surgical practice; develops a successful AI-enabled teaching intervention into a detailed case and publishes it in an institutional or professional teaching-case repository; and presents research on the application of AI in teaching at an academic conference.

2.5.4 Lifelong Learning and Tracking of Advances (CAIP-ME22)

- **Scope and Definition:** Establishes a strong willingness and habit of lifelong learning in the AI era. Develops effective information channels, continuously follows emerging developments, tools, and concepts at the frontier of AI integration with medical education, and proactively incorporates them into a personal professional-development plan.
- **Key Behavioral Indicators:** Regularly reviews selected medical-education technology journals, public communication channels, or professional communities; participates in at least one relevant online or in-person training course, workshop, or academic conference each year; and maintains a continually updated personal collection or notebook of AI teaching tools and resources.

2.5.5 Peer Sharing and Participation in Community Building (CAIP-ME23)

- **Scope and Definition:** Willingly shares personal experience, resources, and insights on AI-enabled teaching at the teaching-and-research group, departmental, or institutional level. Proactively participates in or organizes teaching salons and study groups and helps create a local culture of open collaboration and mutual improvement.
- **Key Behavioral Indicators:** Conducts a brief sharing session in a departmental teaching-and-research activity on a practical topic such as using AI to rapidly produce teaching illustrations; uploads high-quality self-developed AI-generated teaching resources to the institutional sharing platform with instructions for use; and proactively joins or establishes an internal faculty interest group on AI education.

2.5.6 Promotion of an Institutional Culture for AI-Enabled Teaching (CAIP-ME24)

- **Scope and Definition:** Moves beyond the individual classroom and contributes consciously at the departmental or institutional level, through policy recommendations, participation in standard

setting, and service as a trainer, to the development of systems, resource allocation, and a culture that support innovation and dissemination in AI-enabled education.

- Key Behavioral Indicators: When the institution solicits proposals for educational reform, submits policy recommendations to encourage faculty exploration of AI-enabled teaching; participates in drafting institutional Guidelines for the Development and Use of AI Teaching Resources; and serves as an instructor for the introductory AI-teaching module in new-faculty training.

2.5.7 Systems-Level Understanding of the AI Education Ecosystem (CAIP-ME25)

- Scope and Definition: Understands the ecosystem of AI applications in medical education from a broader perspective, including models of industry-academia-research collaboration, the policy and regulatory environment, and global development trends. Has an initial awareness of and perspective on collaboration with industry, other institutions, and international peers.
- Key Behavioral Indicators: Is familiar with the direction of national macro-level policy on medical AI and educational informatization; understands common models and important considerations for collaborating with enterprises on development of teaching tools or research using clinical data; and summarizes several major global trends in the application of AI to medical education, including those in the United Kingdom and the United States.

3. Classification and Hierarchical Implementation Suggestions for CAIP-ME

Based on expert reviews, the competencies outlined in the CAIP-ME framework are categorized into foundational competencies (emphasizing universality, necessity, and scalability) and developmental competencies (highlighting leadership, optional nature, and scenario specificity), as detailed in Table 1.

3.1 Foundational Competencies

A total of 11 foundational competencies are identified: CAIP-ME1–5, CAIP-ME7–8, CAIP-ME15–17, and CAIP-ME19. These competencies are essential for all medical educators, irrespective of their academic background, professional title, or institutional affiliation, to meet the basic requirements of teaching in the AI era, ensuring minimum educational quality and safety standards. They focus on fundamental value recognition, ethical and safety baselines, core concept understanding, compliance with tool operation, and initial applications in teaching and evaluation.

Implementation Notes for Foundational Competencies:

- Baseline and Mandatory Curriculum: These competencies should serve as the "baseline" and "mandatory course" for faculty development at institutions.

- **Training and Assessment:** It is recommended that these competencies be ensured through mandatory, standardized onboarding training and regular assessments. Consider linking them with teacher certification and annual performance evaluations.
- **Key to Success:** Providing systematic, easily accessible training resources and clear operational guidelines to reduce learning and compliance costs for teachers. Technical support and incentives should also be provided to ensure that every teacher meets the required standards.

3.2 Developmental Competencies

There are 14 developmental competencies: CAIP-ME6, CAIP-ME9–14, CAIP-ME18, and CAIP-ME20–25. These competencies are advanced proficiencies that teachers can selectively pursue and deepen based on their subject specialties, course characteristics, personal career development plans, and institutional support conditions. They reflect a teacher's depth of expertise, innovative height, and distinctive development in deeply integrating AI with teaching and research, exploring innovative models, contributing to academic research, and demonstrating leadership.

Implementation Notes for Developmental Competencies:

- **Supportive and Incentivized Environment:** The cultivation of these competencies should not follow a one-size-fits-all approach but instead aim to create a supportive and incentivized environment.
- **Specialized Development Opportunities:** Provide abundant specialized development opportunities such as high-end workshops, interdisciplinary projects, and research grants.
- **Recognition System:** Establish a diversified honor recognition system including teaching innovation awards and micro-certifications for specific skills.
- **Cultural Encouragement:** Foster a team culture that encourages exploration, tolerates failure, and promotes sharing. Institutions should consider achievements in developmental competencies when evaluating teachers for promotions, excellence awards, and resource allocation, thereby stimulating intrinsic motivation and innovative vitality among teachers.

4. Evaluation Methods and Technical Guidance for CAIP-ME

4.1 Evaluation System Design

The purpose of CAIP-ME evaluation is to “promote development” rather than to make simplistic judgments. It is recommended to establish a developmental evaluation system that is multi-dimensional, multi-source, and emphasizes evidence from the teaching process. The evaluation adheres to the following key principles:

1. **Development-oriented:** Evaluation results should first be used to provide teachers with personalized feedback and development suggestions, and only secondarily inform management decisions.
2. **Objectivity:** Emphasis is placed on evidence derived from authentic teaching practices—such as

lesson plans, student work, classroom recordings, platform-generated data, and reflective reports—rather than solely on subjective impressions.

3. Differentiation: Evaluation should be tailored according to teachers’ disciplinary fields (e.g., clinical medicine, basic medical sciences, public health), primary responsibilities (teaching-focused, research-focused, or clinical-focused), and career stages, with appropriate emphasis placed on relevant competencies without requiring proficiency across all areas.

4.2 Evaluation Methods and Evaluators

A comprehensive, consensus-based evaluation approach centered on the “teaching portfolio” is recommended. Objective and quantifiable evaluation indicators should be collected, to the greatest extent possible, through educational information technology platforms and AI Medical Colleges. Evaluation reports should integrate multiple perspectives, as detailed in Table 2.

4.3 Application of Evaluation Results

The fundamental purpose of CAIP-ME-based assessment is to support teacher improvement and professional growth. Specifically, evaluation results can serve the following positive functions:

First, a **personalized development report** can be generated for each participating teacher, systematically diagnosing strengths and areas for growth in their AI teaching competencies and offering concrete, actionable recommendations to clarify their professional development pathway.

Second, evaluation results can serve as a key basis for **professional development certification**. Teachers who demonstrate outstanding performance in specific dimensions may be awarded corresponding micro-credentials or competency certifications—such as “AI-Enhanced Instructional Design Certification” or “Medical AI Ethics Teaching Facilitator”—which can be linked to incentives including continuing education credit accrual or expedited pathways in faculty promotion reviews.

Additionally, evaluation results can inform **institutional resource planning** at medical schools and affiliated teaching hospitals. Institutions can use aggregated findings to identify systemic competency gaps and accordingly allocate resources—such as recruiting specialized personnel, designing targeted training programs, procuring relevant AI tools, or establishing interdisciplinary support teams—thereby enabling precise and efficient investment in faculty development.

5. Implementation and Development Recommendations

5.1 Comprehensive Promotion Strategies at the Institutional Level

To systematically implement this consensus within universities and teaching hospitals, a set of integrated strategies is recommended:

Survey and Planning

Institutions should design assessment tools based on this framework to evaluate the overall level and structural characteristics of teachers' AI literacy. Based on these findings, develop an "Action Plan for Enhancing Teachers' AI Literacy" that aligns with the institution's strategic goals. This plan should outline development objectives, key tasks, resource support, and responsibility assignments over the next 3–5 years.

Layered Piloting and Leading by Example

Avoid a blanket approach. Instead, select one or two departments or teaching units with strong AI application foundations and reform intentions as pilot projects. Focus on core clinical courses to integrate AI deeply into educational reforms. By nurturing “model teachers” and “demonstration cases,” create replicable experiences that can gradually be extended across the institution.

Platform and Resource Construction

Integrate or introduce reliable AI teaching platforms. The education technology center should lead in establishing school-based resource repositories such as “AI Teaching Excellence Case Libraries,” “Tool Usage Guides,” and “Ethical Risk Warning Libraries.” These resources aim to lower teachers’ exploration barriers and reduce trial-and-error costs.

Incentive Mechanisms and Environment Optimization

Incorporate foundational competency standards into basic evaluations and link developmental competency achievements to promotions, awards, and teaching reform fund supports. Actively foster an organizational culture that encourages innovation, tolerates mistakes, and promotes sharing.

By coordinating planning, piloting, support, and incentives, institutions can systematically and steadily enhance their faculty’s overall AI literacy.

5.2 Development Pathways for Individual Teachers

The consensus also provides an actionable framework for individual teacher development:

Self-Assessment and Positioning

Teachers should conduct self-diagnosis using the competency descriptions provided in the framework to determine their current stage—whether it be “cognitive enlightenment,” “tool application,” “deep integration,” or “innovative leadership.” Identify strengths and weaknesses.

Developmental Learning Map

Select 3–5 competencies from the framework that are most urgent and feasible for improvement based on the nature of the courses taught. Develop a detailed annual development plan that includes training

(e.g., enrolling in online courses), practical attempts (e.g., introducing AI tools into a specific teaching segment), and documenting outcomes (e.g., writing reflective journals).

Engagement in Practical Communities

Join internal and external teaching communities, virtual teaching and research offices, or professional societies. Through observation, collaboration, and discussion, achieve mutual experience exchange and capability growth.

5.3 Dynamic Development Mechanism: Continuous Iteration and Co-construction

To maintain the forward-looking nature and vitality of this competency framework, an open, collaborative continuous evolution mechanism must be established:

Regular Revision System

Annually convene expert meetings to systematically review and revise the consensus, considering significant breakthroughs in AI technology, adjustments in medical education policies, empirical research results from large-scale practices, and changes in international standards.

National Practice Case Library and Exchange Network

Encourage institutions to upload excellent practice cases, evaluation tools, and training resources related to the framework onto shared platforms (e.g., China Medical Education MOOC Alliance Platform). Facilitate deep exchanges and collaborative creation through annual forums, thematic workshops, and online salons.

Accompanying Empirical Research

Promote collaboration between educational researchers and frontline teachers to conduct action research on the effectiveness of the framework. Key topics include comparing the effectiveness of different training models, correlating competency development with teaching outcomes, and validating the reliability and validity of evaluation tools. Use research evidence to refine the framework.

Industry Standards and National Recognition

After thorough practical testing and revision, push for recognition by authoritative academic organizations (such as the Chinese Medical Association's Medical Education Branch) and seek attention and endorsement from educational and health administrative authorities. Ultimately, transform the framework into industry guidelines or national standards, guiding the construction of medical teacher teams at broader and higher levels.

6. Conclusion

The framework established in this consensus serves as a dynamic guide for the development of AI literacy among medical educators, aiming to inspire reflection and inform practice. We encourage colleagues across the field to actively engage with the framework in real-world settings, enriching and refining it through collective experience so that it genuinely empowers every teacher.

We call upon educational administrators at all levels, institutional leaders, faculty development centers, and individual teachers to participate proactively—adopting an open mindset, exercising prudent implementation, and engaging in continuous reflection—to drive the practical adoption and iterative evolution of this framework. Together, we can elevate the quality of medical education and cultivate future health professionals who excel in harnessing intelligent technologies, ultimately advancing patient care and public health.

Abbreviations

AI

artificial intelligence

AR

augmented reality

CAIP-ME

Competency Items of Medical Educators' AI Proficiency

CNKI

China National Knowledge Infrastructure

LLM

large language model

VR

virtual reality

Declarations

Secondary Publication and Translation Statement

This article is an authorized English secondary publication of the Chinese-language expert consensus first published in *Acta Academiae Medicinae Sinicae*. The Chinese-language article constitutes the primary publication. The present English version is intended for a different readership and faithfully preserves the substantive consensus content, including the five-dimensional framework, the 25 competency items, their definitions and behavioral indicators, and their classification into 11 foundational and 14 developmental competencies. Permission for this English secondary publication was obtained from the Editorial Office of *Acta Academiae Medicinae Sinicae*, and the relationship

between the two publications was fully disclosed to the editors of BMC Medical Education in the cover letter.¹⁰

Editorial Note for the English Secondary Publication

The substantive consensus text above is a faithful English translation of the primary Chinese publication. For editorial transparency, the evaluation recommendations should not be interpreted as a psychometrically validated assessment instrument. Institutions applying the framework should consider their local policies, resources, faculty roles, educational infrastructure, and governance requirements. Future implementation studies should examine the feasibility, reliability, and educational impact of CAIP-ME in different settings. These clarifications do not revise the five-dimensional framework, the 25 competency items, their scope and behavioral indicators, or their foundational and developmental classification.

Ethics approval and consent to participate

The consensus process involved literature review, expert consultation, and Delphi procedures. It did not involve patients, students, clinical interventions, human biological specimens, patient records, medical images, or identifiable clinical data. The Research Ethics Committee of Guangzhou Women and Children's Medical Center subsequently issued a Certificate of Exemption from Research Ethics Review (Certificate No. 2026-402A01; July 16, 2026). This certificate was a post hoc administrative confirmation obtained after completion of the consensus process rather than prospective ethics approval before expert consultation.

Informed Participation

Experts were informed of the purpose of the consensus process, the voluntary nature of participation, the intended academic use of aggregate findings, and the confidentiality arrangements for their responses. Completion of an expert consultation or Delphi questionnaire was regarded as agreement to contribute professional opinions to the consensus process. Responses used for analysis were de-identified and reported in aggregate.

Consent for publication

Not applicable. The manuscript contains no identifiable individual data, images, or videos.

Funding

This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests

The authors declare that they have no conflicts of interest related to this consensus statement.

Availability of data and materials

Data sharing: No additional participant-level data are publicly available. Aggregate information supporting the consensus is provided in the article and supplementary materials.

Authors' contributions

Mengchun Gong and Hui Pan contributed equally to the conceptualization, consensus design, framework development, and drafting of the manuscript. Jinghui Lu and Zhirong Zeng contributed to the literature review, evidence synthesis, Delphi coordination, data curation, and interpretation. Wei Chen, Hao Wu, Li Luo, Wenwen Sun, Hui Liu, Hang Li, Wei Wang, Yan Luo, and Bohan Zhang contributed to competency development, expert consultation, data collection, interpretation, and critical revision of the manuscript. Xunming Ji supervised the project, provided senior methodological and strategic guidance, and critically revised the manuscript. All authors reviewed and approved the final manuscript and agree to be accountable for the work.

Acknowledgements

Not applicable.

AI-Assisted Language Editing Disclosure

OpenAI Codex was used to assist with English-language editing, consistency checking, and document formatting. No data, references, consensus items, classifications, or substantive conclusions were generated by the tool. The authors reviewed the complete manuscript and take responsibility for its accuracy and integrity.

References

1. Masters K, MacNeil H, Benjamin J, Carver T, Nemethy K, Valanci-Aroesty S, et al. Artificial intelligence in health professions education assessment: AMEE Guide No. 178. *Med Teach*. 2025;47:1410-1424. doi:10.1080/0142159X.2024.2445037.
2. Proumen A, Uribe-Marquez S, Booth GJ, Mitchell JD. Artificial intelligence in medical education. *Anesthesiol Clin*. 2025;43:563-576. doi:10.1016/j.anclin.2025.05.008.
3. 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;9:e48291. doi:10.2196/48291.
4. Ministry of Education Teacher Team Construction Expert Guidance Committee. Guidelines for generative artificial intelligence application for teachers (version 1) (in Chinese). *China Education Daily*. 2025 Dec 28. Available from: <https://mp.weixin.qq.com/s/do2SwTai9xyxnNG9efQ13Q>. [Last accessed on 2026 Aug 30].

5. Zhang J. Are we ready for AI-augmented generalists? *BMJ Evid Based Med*. 2026;31:139-140. doi:10.1136/bmjebm-2024-113597.

6. Saroha S. Artificial intelligence in medical education: promise, pitfalls, and practical pathways. *Adv Med Educ Pract*. 2025;16:1039-1046. doi:10.2147/AMEP.S523255.

7. Weidener L, Fischer M. Proposing a principle-based approach for teaching AI ethics in medical education. *JMIR Med Educ*. 2024;10:e55368. doi:10.2196/55368.

8. Gong MC, Ma YH, Pan H, Bai H, Dai H, Chen W, et al. Expert consensus on ethical governance for clinical applications of generative medical artificial intelligence (2025 edition) (in Chinese). *Acta Acad Med Sin*. 2026;48:24-37. doi:10.3881/j.issn.1000-503X.17220.

9. Gong MC, Pan H, Liu H, Tong K, Li J, Ma YH, et al. Expert consensus on the artificial intelligence proficiency competency list and assessment framework for medical students (2025 edition) (in Chinese). *Acta Acad Med Sin*. 2026;48:13-23. doi:10.3881/j.issn.1000-503X.17210.

10. Pan H, Gong MC, Lu JH, Zeng ZR, Chen W, Wu H, et al. Expert consensus on artificial intelligence proficiency for medical educators (2025 edition) (in Chinese). *Acta Acad Med Sin*. 2026;48:1-12. doi:10.3881/j.issn.1000-503X.17190.

Tables

Table 1. Classification of the 25 CAIP-ME competencies

Classification	Number	CAIP-ME Content
Foundational Competencies (F)	CAIP-ME1	Recognition of the Value of AI-Enabled Medical Education
	CAIP-ME2	Identification and Assessment of Ethical Risks in Medical AI
	CAIP-ME3	Guidance on Academic Integrity and Standards for AI Use
	CAIP-ME4	Ability to Promote Equity and Justice in Educational Practice
	CAIP-ME5	Understanding of Core AI Concepts and Medical Applications
	CAIP-ME7	Operational Competency in Data Security and Privacy Protection
	CAIP-ME8	AI-Assisted Literature Retrieval and Analysis
	CAIP-ME15	Design of AI-Supported Formative Assessment
	CAIP-ME16	Interpretation of Learning Data and Instructional Intervention
	CAIP-ME17	Diagnosis and Support of Students' Digital Literacy
Developmental Competencies (D)	CAIP-ME19	Reflection on and Optimization of AI-Enabled Teaching Practice
	CAIP-ME6	Evaluation and Selection of Medical AI Tools
	CAIP-ME9	Basic Processing and Visualization of Research Data
	CAIP-ME10	AI-Enhanced Instructional Design
	CAIP-ME11	Application of Intelligent Virtual Simulation in Teaching
	CAIP-ME12	AI-Assisted Creation of Teaching Resources
	CAIP-ME13	Organization of Human-AI Collaborative Classroom Teaching
	CAIP-ME14	Integration of Clinical AI Tools into Teaching Scenarios
	CAIP-ME18	Tracking and Evaluation of Students' AI Proficiency Development

Classification	Number	CAIP-ME Content
	CAIP-ME20	Interdisciplinary Collaboration and Communication
	CAIP-ME21	Research on AI-Enabled Educational Reform and Translation of Outcomes
	CAIP-ME22	Lifelong Learning and Tracking of Advances
	CAIP-ME23	Peer Sharing and Participation in Community Building
	CAIP-ME24	Promotion of an Institutional Culture for AI-Enabled Teaching
	CAIP-ME25	Systems-Level Understanding of the AI Education Ecosystem

Table 2. Evaluation suggestions for AI proficiency of medical educators

Evaluation Method	Description	Evaluation Subject
Evaluation of teaching practice evidence	Teachers systematically collect and submit materials reflecting their AI teaching competencies, such as lesson plans, courseware, student achievement analysis reports, teaching reflections, innovative cases, etc	Instructional development specialist, peer expert, AI technology consultant
Classroom/clinical teaching observation	Observe teachers' behavioral manifestations in using AI for teaching or mentoring in natural settings through on-site observation or video recording	Teaching supervisor, peer teacher
Student feedback survey	Through structured questionnaires or in-depth interviews, collect students' perceptions of the effectiveness, fairness, and guiding power of AI teaching applications used by teachers	Student
Platform data tracking	With the authorization of teachers, anonymous and aggregated behavioral and performance data are collected from teaching platforms and assessment systems as supporting evidence	Educational Technology Center, data analyst
Self-assessment and reflection	Teachers conduct self-assessment against the framework and write job reports or development plans	The educator

Figures

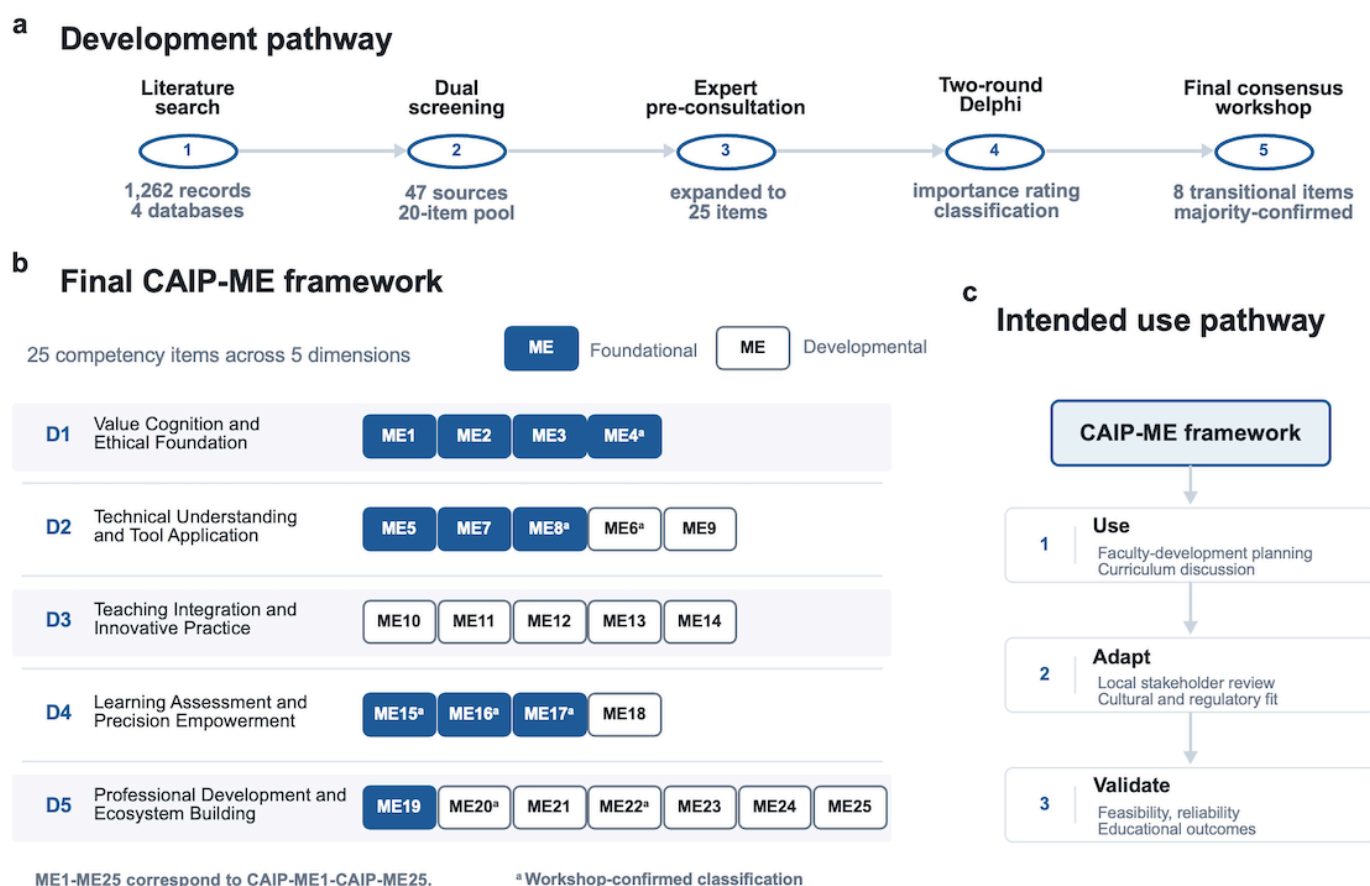


Figure 1

Development pathway, final structure, and intended use of the CAIP-ME framework. Panel a summarizes literature screening, expert pre-consultation, the two-round Delphi process, and final consensus workshop. Panel b presents the five dimensions and 25 competency items, including foundational and developmental classifications. Panel c summarizes the intended use, contextual adaptation, and validation pathway.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile2Searchstrategies.docx](#)
- [Additionalfile5Firstroundmeetingminutes.docx](#)
- [Additionalfile12Glossary.docx](#)
- [Additionalfile13CHERRIESchecklist.docx](#)
- [Additionalfile7Responsetoexpertreview.docx](#)
- [Additionalfile6Firstroundexpertcomments.docx](#)

- [Additionalfile11DetailedCAIPMEcompetencytable.docx](#)
- [Additionalfile9Finalizationexpertcomments.docx](#)
- [Additionalfile10DetailedDelphiresultsandquestionnaires.docx](#)
- [Additionalfile3PRISMA2020flowdiagram.docx](#)
- [Additionalfile8Finalizationmeetingminutes.pdf](#)
- [Additionalfile4Expertpanelcharacteristics.docx](#)
- [Additionalfile1RawPRISMAflowdata.xlsx](#)
- [Appendix.docx](#)