Measuring AI Literacy Among Medical Students: An Item Set*

Weiguang Wang[1,†], Shuangqing Zhai[1,*,†]

[1] Beijing University of Chinese Medicine, No. 11 Bei San Huan Dong Lu, Chaoyang District, Beijing, 100029, China

Abstract

The rapid development of artificial intelligence (AI) has brought revolutionary changes to medicine, and cultivating the AI literacy of medical students has become an important part of medical education. How to measure and analyze the artificial intelligence literacy of medical students is an important part of it. This study aims to analyze the evaluation system of medical students' artificial intelligence literacy through a systematic and extensive literature review, and propose future research directions. We have comprehensively evaluated the research papers on dispersed artificial intelligence literacy based on literature review, At the same time, it is proposed to optimize the existing assessment framework from five dimensions: consciousness, knowledge, thinking, skills, and ethics. Our measurement framework is more in line with the characteristics of medical students and has good application prospects, making important contributions to AI driven medical work.

Keywords

review, AI literacy, medical college students, assessment framework, learning Analytics

1. Introduction

With the increasing integration of artificial intelligence into key medical domains—including disease diagnosis, medical image analysis, treatment planning, and health management—the traditional medical paradigm is undergoing a profound transformation. Future physicians will be required to possess not only foundational medical knowledge and clinical skills but also to develop competencies in understanding, critically evaluating, and effectively utilizing AI-based tools. Consequently, fostering and assessing AI literacy among medical students has emerged as a central focus of global medical education reform.

A growing body of research has examined the concept of AI literacy among students. Kandlhofer(2016)[1] initially introduced the notion of AI literacy, which was further developed and refined by Kong(2021)[2]. In a broad sense, AI literacy refers to an individual's comprehensive development within environments mediated by AI technologies. For instance, Laupichler(2023)[3] proposed an evaluative framework encompassing 38 items related to AI competencies. More recently, Pinski(2024)[4] and Ng(2024)[5] contributed to the discourse by proposing AI literacy assessment models from differing perspectives.

AI literacy among medical students constitutes a distinct and increasingly important subset of general student AI literacy. Karaca(2021)[6] developed an AI literacy scale specifically tailored to medical students (MAIRS-MS). Laupichler(2024)[7] empirically investigated AI literacy levels within this population, while Busch(2024)[8] explored the integration of AI applications into medical education curricula. Additionally, Naseer(2025)[9] examined strategies for embedding AI content into medical training programs.

* Joint Proceedings of LAK 2026 Workshops, co-located with the 16th International Conference on Learning Analytics and Knowledge (LAK 2026), Bergen, Norway, April 27 – May 1, 2026

1,* Corresponding author.

† These authors contributed equally.

202001031@bucm.edu.cn (W. Wang); zsq2098@163.com (S. Zhai)

0009-0004-0786-4167 (W. Wang); 0000-0002-7226-3577 (S. Zhai)

Recent scholarship has increasingly emphasized the significance of generative AI (GenAI) literacy
. Bozkurt(2024)[10], for example, argued that a GenAI literacy framework must highlight the ethical distinctions between generative and traditional AI systems. Similarly, Annapureddy(2025)[11] proposed a literacy framework specifically tailored to generative AI contexts.

Despite these advancements, a unified and systematic standard for assessing the AI literacy of medical students—particularly in relation to generative AI—remains lacking. Moreover, existing assessment instruments often emphasize technical competencies while inadequately integrating authentic medical practice scenarios. In response to this gap, the present study aims to develop a multidimensional and comprehensive evaluation framework. This endeavor seeks to contribute to the scientific and standardized advancement of AI literacy education among medical students.

2. Modifications

To ensure the scientific validity and effectiveness of the assessment framework, the design of assessment items in this study adhered to the principles of systematicity, professionalism, and feasibility. The framework was progressively constructed using the following specific methods.

2.1. Systematic Literature Review and Policy Analysis

Following extensive searches of Chinese and English databases, a systematic review of domestic and international assessment tools and core competency requirements for digital literacy, health informatics, and medical AI education was performed. Concurrently, authoritative documents—including those from the World Federation for Medical Education (WFME) and China's "Standards for Medical Education"—were analyzed to extract information and technology competency requirements for medical students. AI-related core elements were extracted from these sources to form the preliminary item pool.

2.2. Expert Interview Method

The interviewees primarily included medical education experts, clinical physicians, medical informatics researchers, medical ethics teachers, and students. Three individuals were selected for each role. By conducting interviews with these stakeholders, the initial item pool was revised and refined, leading to the selection of items with high consensus and strong representativeness.

3. RESULTS

Based on the aforementioned design and validation process, an evaluation framework comprising 5 dimensions and 18 items has been developed, as shown in the Table 1.

Table 1
Medical Students' AI Literacy Assessment Items

Dimension	Item
Awareness	1. Awareness of Development Trends
	2. Awareness of Using Tools
	3. Awareness of Exploring AI Capabilities
	4. Awareness of Integrating AI Technology with Medical Theory
Knowledge	5. Basic Concepts

	6. Core Principles
	7. Data Fundamental
Skills	8. Proficiency in Tool Operation
	9. Critical Interpretation and Evaluation of Result
	10. Foundational Understanding and Handling of Data Structures
Ethics	11. Data Security
	12. Privacy Protection
	13. Intellectual Property
	14. Human-AI Relationship

3.1. Awareness Dimension

The Awareness dimension refers to students' conscious awareness of AI use and exploration, including four key aspects: Awareness of AI Development Trends, Awareness of Using AI Tools, Awareness of Exploring AI Capabilities, and Awareness of Integrating AI Technology with Medical Theory.

Awareness of AI Development Trends requires students to grasp AI, its medical development trends, internal logic, and potential directions from a dynamic, forward-looking perspective—beyond merely acknowledging its rapid progress—and propose developmental ideas.

Awareness of Using AI Tools denotes students' spontaneous inclination to seek" AI assistance" in learning, creation, or problem-solving scenarios, translating this awareness into action and evolving it
into an internalized thinking habit and work methodology beyond mere tool recognition.

Awareness of Exploring AI Capabilities is a curiosity-driven, practice-oriented awareness urging students not to settle for AI as a "black box" for predefined tasks but to actively explore its capability
boundaries and operating principles.

Awareness of Integrating AI Technology with Medical Theory requires moving beyond viewing AI as a mere tool to establish a bidirectional integration paradigm: using AI to uncover complex correlations while firmly relying on medical theory for causal inference, turning AI from a supportive tool into a dynamic driver of theoretical progress.

3.2. Awareness Dimension

The Knowledge dimension refers to student' understanding of AI-related knowledge, encompassing Basic Concepts, Core Principles, and Data Fundamentals.

Basic Concepts serve as the foundation for AI understanding, enabling students to accurately identify and articulate core elements. Students should grasp the definitions and core connotations of key terms as well as AI lifecycle-related process concepts.

Core Principles elucidate AI's operational logic without requiring complex mathematical depth: machine learning learns data patterns rather than relying on hard-coded rules; neural networks

extract features through layered information processing; large language models generate coherent text via probability prediction; and generative AI "hallucinations" are fundamentally irreducible.

Data Fundamentals reflect the core of AI systems, as understanding data equates to grasping the origin of AI behavior. Students must acknowledge that: AI performance is highly dependent on the scale, quality, and representativeness of training data; data biases or blind spots directly result in biased or discriminatory model outputs; and data collection and usage entail significant privacy and ethical implications.

3.3. Skills Dimension

The Skills dimension denotes students' AI application competence, encompassing Proficiency in Tool Operation, Critical Interpretation and Evaluation of Results, and Foundational Understanding and Handling of Data Structures.

Proficiency in Tool Operation is the practical capacity to use specific AI tools effectively, focusing on mastering general operational logic and core functions: efficiently decomposing tasks, designing clear structured "prompts" for target content; selecting/switching tools per task features, integrating multiple AI applications flexibly; and using basic troubleshooting/optimization via iterative prompting or parameter adjustments.

Critical Interpretation and Evaluation of Results involves sound judgment on AI outputs to use them effectively: identifying their reliability and limitations by cross-verifying information for "hallucinations" or factual errors; assessing logical consistency, relevance, and practicality of AI-generated content; and discerning implicit biases or security risks.

Foundational Understanding and Handling of Data Structures is the basic ability to prepare, organize, and manage data for effective AI use—dependent heavily on input data quality. It includes understanding task-specific input format requirements; conducting simple data cleaning/structured organization with preliminary data awareness; and grasping basic data privacy/security principles to handle sensitive information as needed.

3.4. Ethics Dimension

The Ethics dimension refers to students' awareness of AI ethical issues, including Data Security, Privacy Protection, Intellectual Property, and Human-AI Relationship.

Data Security protects AI-lifecycle data from leakage, tampering, loss, or misuse. Students should know AI systems are vulnerable (models may error or leak training data if attacked), avoid sensitive internal data in AI tools, and check platforms' data security commitments.

Privacy Protection means respecting and legally using personal info. Students should know their rights and be cautious about AI using their data—especially high-risk biometrics.

Intellectual Property covers AI-generated content copyright, legal training data sources, and lines between human creation and AI assistance.

The Human-AI Relationship involves rights, responsibilities, and social impact in human-AI interaction. AI is a tool—humans bear responsibility for AI decisions. Socially, guard against algorithmic bias worsening inequities and over-reliance weakening human autonomy.

4. CONCLUSION

This study aims to analyze the evaluation system of medical students' AI literacy through a systematic and extensive literature review, and propose future research directions. We have comprehensively evaluated the research papers on dispersed artificial intelligence literacy based on literature review. It is proposed to optimize the existing assessment framework from 4 dimensions and 14 items. The dimensions are consciousness, knowledge, skills, and ethics. Our measurement framework is more in line with the characteristics of medical students and has good application prospects, making important contributions to AI driven medical work.

5. Acknowledgments

This work was supported by National Traditional Chinese Medicine Education Examination Project(Number: TE2024004).

Declaration on Generative AI

During the writing of this manuscript, AI technology (DeepSeek) was used solely for grammar, spelling, and formatting assistance.

References

[1] M. Kandlhofer, G. Steinbauer, S. Hirschmugl-Gaisch, P. Huber, Artificial intelligence and computer science in education: From kindergarten to university, in: 2016 IEEE Frontiers in Education Conference (FIE), IEEE, 2016, pp. 1–9. doi:10.1109/FIE.2016.7757570.

[2] S.-C. Kong, W. Man-Yin Cheung, G. Zhang, Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds, Computers and Education: Artificial intelligence. 2 (2021) 100026. doi:10.1016/j.caeai.2021.100026.

[3] M. C. Laupichler, A. Aster, T. Raupach, Delphi study for the development and preliminary validation of an item set for the assessment of non-experts' AI literacy, Computers and Education: Artificial intelligence. 4 (2023) 100126. doi:10.1016/j.caeai.2023.100126.

[4] M. Pinski, A. Benlian, AI literacy for users – A comprehensive review and future research directions of learning methods, components, and effects, Computers in Human Behavior: Artificial Humans. 2 (2024) 100062. doi:10.1016/j.chbah.2024.100062. arXiv:7.

[5] D. T. K. Ng, J. K. L. Leung, K. W. S. Chu, M. S. Qiao, AI Literacy: Definition, Teaching, Evaluation and Ethical Issues, Proceedings of the Association for Information Science and Technology. 58 (2024) 504–509. doi:10.1002/pra2.487.

[6] O. Karaca, S. A. Çalışkan, K. Demir, Medical artificial intelligence readiness scale for medical students (MAIRS-MS) – development, validity and reliability study, BMC Medical Education. 21 (2021) 112. doi:10.1186/s12909-021-02546-6.

[7] M. C. Laupichler, A. Aster, M. Meyerheim, T. Raupach, M. Mergen, Medical students' AI literacy and attitudes towards AI: A cross-sectional two-center study using pre-validated assessment instruments, BMC Medical Education. 24 (2024) 401. doi:10.1186/s12909-024-05400-7.

[8] F. Busch, L. Hoffmann, D. Truhn, E. Ortiz-Prado, M. R. Makowski, K. K. Bressem, L. C. Adams, COMFORT Consortium, Global cross-sectional student survey on AI in medical, dental, and veterinary education and practice at 192 faculties, BMC Medical Education. 24 (2024) 1066. doi:10.1186/s12909-024-06035-4.

[9] M. A. Naseer, S. Saeed, A. Afzal, S. Ali, M. G. R. Malik, Navigating the integration of artificial intelligence in the medical education curriculum: A mixed-methods study exploring the perspectives of medical students and faculty in Pakistan, BMC Medical Education. 25 (2025) 273. doi:10.1186/s12909-024-06552-2.

[10] A. Bozkurt, Why Generative AI Literacy, Why Now and Why it Matters in the Educational Landscape? Kings, Queens and GenAI Dragons, International Council for Open and Distance Education. 16 (2024) 283–290. doi:10.55982/openpraxis.16.3.739.

[11] R. Annapureddy, A. Fornaroli, D. Gatica-Perez, Generative AI Literacy: Twelve Defining Competencies, Association for Computing Machinery. 6 (2025) 1–21. doi:10.1145/3685680.arXiv:2412.12107.