Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and validation of the artificial intelligence literacy scale for medical students (AILS-MS)

Weijia Feng
Chongqing Medical University
Yaoyue Hu
Chongqing Medical University
Jialei Tang
Chongqing Medical University
Han Zhou
Chongqing Medical University
Hong Xu
xuhong@cqmu.edu.cn

Chongqing Medical University

Research Article

Keywords: Artificial intelligence literacy, Medical students, Scale development, Reliability and validity testing

Posted Date: September 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10632038/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

Background

With the rapid expansion of artificial intelligence (AI) applications in healthcare, wellness, and education, there is an urgent need to assess and enhance AI literacy among future healthcare professionals. This study aimed to develop and validate an AI literacy scale for medical students.

Methods

Guided by UNESCO's *AI Competency Framework for Students*, an initial scale was constructed through literature review, semi-structured interviews, and two rounds of Delphi expert consultation. The scale was subsequently revised and validated using item analysis, exploratory factor analysis, confirmatory factor analysis, and reliability testing.

Results

The model demonstrated good fit (χ^2/df = 2.171; CFI = 0.919; TLI = 0.905; RMSEA = 0.063). The final instrument comprises 20 items organized into four dimensions: technology and application, humanistic values, ethics and fairness, and AI-enhanced learning.

Conclusion

This study integrates core medical education requirements into a globally informed AI competency framework and develops a specialized instrument for assessing AI literacy among medical students. The resulting scale provides a quantitative foundation to support competency-based assessment, curriculum reform, and institutional planning as AI becomes increasingly embedded in medical education.

Introduction

Artificial intelligence (AI) has evolved at an unprecedented pace in recent years, turning the vision of human-AI collaborative diagnostics into a tangible reality [1]. The application of AI in healthcare, wellness, and education has grown explosively, achieving routine deployment across core clinical domains such as medical imaging interpretation, clinical risk prediction, and drug discovery [2, 3], thereby profoundly reshaping the delivery of healthcare services. Against this backdrop, AI literacy has become an essential competency for contemporary healthcare professionals. Without it, clinicians may remain mired in occupational burnout due to their inability to delegate repetitive tasks to AI systems, limiting their capacity to manage complex clinical decision-making and provide humanistic care. Moreover, a lack of understanding of algorithmic principles and uncritical reliance on AI outputs may lead to diagnostic errors that compromise patient safety and provoke ethical controversies [4].

Therefore, AI literacy serves not only as a tool for improving efficiency but also as an essential safeguard for clinical quality, patient safety, and sustainable professional practice.

In response to this transformation, medical education institutions worldwide have increasingly incorporated AI literacy into their core competency frameworks. In June 2024, the European Union enacted the world's first Artificial Intelligence Act and explicitly called for the incorporation of AI literacy into educational systems. Three months later, the United Nations Educational, Scientific and Cultural Organization (UNESCO) released the *AI Competency Framework for Students* (hereinafter referred to as "*the UNESCO framework*"), further delineating key dimensions of AI literacy, including human-centered thinking, AI ethics, and AI system design [5]. Building on these global initiatives, Chinese scholars in February 2026 developed an expert consensus titled *Consensus on the Competency Inventory and Assessment Framework of Artificial Intelligence Literacy for Medical Students* (hereinafter "*the Consensus*") [6]. This consensus establishes a localized framework comprising 21 specific indicators, with particular emphasis on applications of generative artificial intelligence (GenAI) and alignment with China's healthcare regulatory context. The sequential issuance of these documents has substantially clarified the content, objectives, and pedagogical approaches for cultivating AI literacy in medical education.

A systematic review of existing assessment instruments reveals several critical gaps. Current research in China remains largely confined to the development of conceptual frameworks and expert consensus statements, with insufficient large-scale empirical data to support rigorous validation of structural validity and reliability. Furthermore, most existing scales predate the widespread adoption of GenAI and therefore lack coverage of key emerging competencies, such as prompt engineering, awareness of algorithmic bias, and hallucination detection [7–9]. More critically, no standardized AI literacy scale specifically designed for medical students exists in China. Available tools are primarily tailored for general undergraduate populations and thus fail to incorporate healthcare-specific contexts, including clinical data privacy and patient safety considerations [10, 11]. Although specialized instruments for medical students exist internationally, these were developed within Western healthcare systems and legal frameworks, raising significant concerns regarding cultural and contextual transferability if applied directly in the Chinese setting [12, 13]. Therefore, there is an urgent need in Chinese medical education to develop a psychometrically robust assessment tool that aligns with international standards, demonstrates empirical validity, and integrates both medical specificity and local contextual relevance.

Accordingly, this study strictly adhered to established scale development guidelines. Grounded in the UNESCO *AI Competency Framework for Students* as its theoretical foundation, an initial item pool was constructed through a multi-phase process involving comprehensive literature review, semi-structured interviews with key stakeholders, and multiple rounds of Delphi expert consultation. Subsequently, empirical data were collected and subjected to rigorous psychometric analyses—including exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and reliability testing to develop and validate a standardized instrument for assessing AI literacy among medical students in China. This endeavor aims to address the current gap in domestically developed assessment tools in this domain and to provide

medical schools with a scientifically sound and reliable quantitative basis for evaluating student competencies, informing curriculum reform, and guiding institutional policy development.

Methods

2.1. Participants

A cross-sectional survey was conducted between November and December 2025 among medical students from three medical universities in Chongqing and Hunan Province. The sample was split by study site to support the psychometric validation of the measurement instrument. Specifically, students from CQMU constituted the exploratory factor analysis (EFA) sample, while the confirmatory factor analysis (CFA) sample included students from Changsha Medical University and Hunan University of Medicine. All participants were aged 18 years or older, majoring in clinical medicine or related health disciplines, and provided informed consent prior to participation. The study protocol was approved by the Ethics Committee of Chongqing Medical University (No.2026 – 145)

2.2. Data collection

Data were collected via a self-administered, anonymous online questionnaire distributed through official university learning management systems or class WeChat/QQ groups during scheduled class hours to maximize response rates and minimize selection bias. The questionnaire assessed sociodemographic and behavioral variables, including gender, age, education level, and frequency of AI use, along with the self-developed Artificial Intelligence Literacy Scale for Medical Students (AILS-MS). The full scale and its scoring instructions are provided in Supplementary file 1. Prior to full-scale deployment, the questionnaire underwent pilot testing (n = 30) to ensure clarity and technical functionality.

2.3. Statistical Analyses

This study employed EFA with principal component extraction and varimax rotation to examine the underlying factor structure and assess construct validity. The number of factors to retain was determined based on eigenvalues greater than 1.0, inspection of the scree plot, and with a value of Kaiser-Meyer-Olkin of 0.6 and above. Items with factor loadings below 0.40 were considered for removal. Subsequently, CFA was performed to evaluate the model fit of the factor structure, with the extracted factors specified as latent variables and the corresponding items as observed indicators. The internal consistency of the overall scale was evaluated by Cronbach's alpha. Statistical analyses were performed using IBM SPSS 26.0 and AMOS 24.0.

Results

A total of 558 and 310 questionnaires were received in the EFA and CFA phases, respectively; of these, 524 and 296 were valid after excluding responses exhibiting evident patterns indicative of invalid responding. The demographic characteristics of the participants are summarized in Table 1.

Table 1
Demographic characteristics of participants

Characteristics	EFA (*n* = 524)		CFA (*n* = 296)		Total (*n* = 820)	
	N	%	N	%	N	%
Gender						
Male	231	44.1	137	46.3	368	44.9
Female	293	55.9	159	53.7	452	55.1
Age						
18–22	332	63.4	190	64.2	522	63.7
≥ 23	192	36.6	106	35.8	298	36.3
Education level						
Undergraduate	411	78.4	192	64.9	603	73.5
Graduate or above	113	21.6	104	35.1	217	26.5
AI usage frequency						
Daily	210	40.1	142	48.0	352	42.9
Several times a week	179	34.2	119	40.2	298	36.3
Several times a month	79	15.1	23	7.8	102	12.4
Rarely	53	10.1	11	3.7	64	7.8
Never	3	0.6	1	0.3	4	0.5

3.1. Item pool generation

A systematic literature search was conducted to identify studies on AI literacy in medical education contexts. Key search terms included “artificial intelligence literacy,” “AI literacy,” “medical students,” “scale development,” and “competency framework.” Retrieved articles were subjected to content analysis to systematically synthesize the core dimensions of AI literacy reported in the existing literature (Table 2). Building on the UNESCO *AI Competency Framework for Students* and accounting for the unique contextual demands of medical education, we adapted the original framework’s dimensions to formulate the primary indicators of the present scale.

Table 2
Existing indicators of AI literacy among college students

Research Literature	First-level Indicators	Second-level Indicators	Number of Items	Year
The Development and Validation of the Artificial Intelligence Literacy Scale for Chinese College Students (AILS-CCS) (Ma et al., 2024) [14]	Awareness, Usage, Evaluation, Ethics	None	15	2024
An Examination of AI Literacy Among University Students: Current Status and Multifaceted Influencing Factors (Zhou et al., 2024)[8]	AI knowledge, AI skills, AI attitudes and values, AI ethics	None	25	2024
Developing a validated assessment of artificial intelligence literacy for Chinese university students based on educational objectives taxonomy (Han et al., 2025)[15]	AI knowledge, AI application, AI attitude, AI ethics, and AI innovation	None	63	2025
Construction of AI Literacy Evaluation System for College Students and an Empirical Study at Wuhan University (Wu et al., 2025)[16]	AI attitude, AI knowledge, AI capability, AI ethics	10	25	2025
Development of an Artificial Intelligence Literacy Indicator System for University Students Using a Combined AHP-BPNN Approach (Ding et al., 2025)[10]	Knowledge and Comprehension, Skills and Application, Evaluation and Creation, Ethics and Morality	17	17	2025
Construction of AI Literacy Evaluation System for College Students from the Continuum Perspective (Li et al., 2025)[11]	Awareness and Attitudes, Knowledge and Skills, Higher-Order Thinking, Intelligence Ethic, Value Realization	19	19	2025
Development and Application of the Assessment Tool for College Students'	GenAI Awareness Awakening, GenAI Knowledge Construction, GenAI Skills Development,	13	38	2025

Research Literature	First-level Indicators	Second-level Indicators	Number of Items	Year
Generative Artificial Intelligence Literacy (Liu et al., 2025)[17]	GenAI Mindset Cultivation, GenAI Ethical Cognition			
Development and validation of Generative AI Competence Scale (GenAI Comp) among university students (Lee et al., 2025)[18]	Information and Data Literacy, Communication and Collaboration, Digital Content Creation, Safety and Ethics, Problem Solving	None	22	2026
Expert Consensus on the Artificial Intelligence Proficiency Competency List and Assessment Framework for Medical Students (2025 Edition) (Gong et al., 2026) [6]	Knowledge, Skill, Attitude	None	21	2026

To enhance the instrument's relevance and professional alignment within medical education settings, semi-structured interviews were conducted. The semi-structured interview guide was developed for this study based on the objectives and a review of relevant literature. The interview guide is provided as Supplementary file 2. Participants were selected through maximum variation sampling to represent diverse academic backgrounds such as clinical medicine, preventive medicine, and intelligent medical engineering from institutions in Chongqing, with varying levels of AI exposure. Data collection continued until thematic saturation was achieved, defined as no new thematic categories emerging after three consecutive interviews, resulting in a total of 20 interviews.

Thematic analysis was performed using NVivo 12.0 Plus. Two researchers independently conducted initial coding of the transcripts following a three-stage process of labeling, conceptualization, and categorization to develop a preliminary coding framework. Subsequently, two additional researchers independently reviewed the coded data to establish intercoder reliability. During the categorization phase, initial categories were further refined through secondary clustering that integrated core values of medical education such as patient-centeredness, humanistic care, and social equity to enhance the content validity of the final indicators. The resulting qualitative findings were used to supplement and semantically refine the initial item pool, yielding a total of 30 initial items, which were further synthesized into four core themes.

3.2. Item screening and refinement

A Delphi study was conducted to refine the initial item pool through multiple rounds of expert consultation [19]. Twenty experts from medical universities and comprehensive tertiary hospitals in

Chongqing participated in the process. Expert selection emphasized interdisciplinary representation and practical relevance, covering fields such as medical education, intelligent medical engineering, clinical medicine, and medical ethics. The panel comprised frontline professionals with extensive experience in clinical practice, teaching, or administration. Both rounds achieved a 100% response rate, suggesting high levels of expert participation. Consensus was assessed using the coefficient of variation (CV). An item was retained only if it met both criteria: a mean importance rating of at least 4.0 on a 5-point scale and a CV ≤ 0.25[20].

In the first round, seven items failed to meet these thresholds and were consequently removed. The revised instrument incorporating this feedback formed the basis for the second-round questionnaire. In the second round, respondents suggested merging redundant items, simplifying overly technical phrasing, and adding concrete examples to improve clarity. In response, we conducted a thorough review to clarify semantic ambiguities and conceptual overlaps, refining wording. All remaining items in second round achieved mean importance scores greater than 4.0 and CV below 0.25, indicating strong consensus among panelists. Given that all items met the a priori retention criteria, a third round was deemed unnecessary.

Following the Delphi process, cognitive debriefing interviews were conducted with five medical students. In response to their feedback on clarity, semantic precision, and readability, minor linguistic revisions were made. The final version of the scale, comprising 22 items, demonstrated adequate face and content validity, and was prepared for EFA.

3.3. Exploratory factor analysis (EFA)

To preliminarily explore the underlying factor structure of the scale, we conducted EFA on a pilot sample of 524 respondents. Prior to the main analysis, data suitability for factor analysis was assessed. The Kaiser–Meyer–Olkin measure of sampling adequacy was 0.911, and Bartlett's test of sphericity was significant, $\chi^2 = 5082.698$, $p < 0.001$. Principal component analysis with varimax rotation was employed for factor extraction. The number of factors to retain was determined based on eigenvalues greater than 1, inspection of the scree plot, and the interpretability of the resulting factor structure. During item refinement, items with factor loadings below 0.40 were considered non-significant and removed. Additionally, items that loaded on two or more factors with differences in loadings less than 0.10 were also excluded. Exploratory factor analysis yielded four factors with eigenvalues greater than 1, accounting for 62.057% of the total variance. The rotated factor matrix is presented in Table 3.

Table 3
Rotated factor matrix

Factors	1	2	3	4
Technology and application				
I understand the fundamental workflow of artificial intelligence (AI), including its key components such as learning, reasoning, decision-making, and execution.	**0.842**	0.156	0.121	0.086
I actively follow the latest developments and advancements of AI applications in the medical field.	**0.727**	0.089	0.281	0.192
I am proficient in using effective prompts to guide AI systems toward generating desired outputs.	**0.828**	0.085	0.120	0.130
When confronted with complex clinical or academic challenges, I can leverage AI to devise innovative solutions.	**0.644**	0.274	0.272	0.238
I am capable of developing a novel AI tool (e.g., constructing a new AI agent).	**0.741**	-0.020	0.038	0.132
Humanistic values				
I consciously avoid overreliance on AI and maintain independent clinical reasoning skills.	0.084	**0.754**	0.180	0.160
I regard AI as an auxiliary tool, with final decision-making authority residing solely with healthcare professionals.	0.049	**0.739**	0.136	0.086
I am able to critically evaluate the outputs generated by AI systems.	0.089	**0.639**	0.333	0.277
I can verify the accuracy and reliability of AI-generated content through multiple approaches (e.g., consulting authoritative literature, tracing data sources, and performing logical validation).	0.156	**0.595**	0.335	0.225
I am aware of the potential adverse societal and interpersonal impacts that AI may entail.	-0.387	**0.426**	0.175	0.189
I recognize that AI can optimize the allocation of healthcare resources (e.g., enhancing equitable access for underserved or remote populations).	0.173	**0.664**	0.161	0.221
Ethics and fairness				
I am familiar with ethical guidelines and relevant healthcare regulations governing AI use.	0.048	0.245	**0.734**	0.221
I can avoid AI usage practices that violate ethical standards.	0.061	0.265	**0.620**	0.127
I implement measures to safeguard data security and privacy when utilizing AI technologies.	0.147	0.225	**0.692**	0.293

Factors	1	2	3	4
Technology and application				
I understand that AI systems may exhibit algorithmic bias (e.g., the quality and origin of training data can influence AI outputs).	0.182	0.358	**0.624**	0.192
I am able to mitigate the effects of algorithmic bias by incorporating more diverse training datasets.	0.329	0.066	**0.711**	0.181
AI-enhanced learning				
I can assess whether AI-generated outputs align with the requirements of specific contexts or scenarios.	0.193	0.318	0.201	**0.766**
I am capable of refining and optimizing AI-generated content according to practical needs.	0.134	0.225	0.240	**0.781**
I can apply AI in interdisciplinary settings (e.g., using AI to support clinical diagnosis).	0.173	0.143	0.232	**0.767**
I view AI as a vital tool supporting my commitment to lifelong learning.	0.143	0.199	0.198	**0.708**
Eigenvalue	3.384	3.141	3.001	2.885
Variance explained (%)	16.918	15.706	15.007	14.426
Cumulative variance explained (%)	16.918	32.624	47.631	62.057

Bold: Highest factor loading.

3.4. Confirmatory factor analysis (CFA)

To further examine the construct validity of the scale, CFA was conducted using AMOS 24.0 based on 296 formal test samples. As shown in Table 4, the initial model exhibited inadequate fit, with both the CFI and TLI slightly below the Acceptable value of 0.90, suggesting that the hypothesized structure required refinement [21]. To improve model fit, three correlated error sources with both statistical and contextual positive correlations were added to the model, following a comprehensive evaluation of the measurement objectives and item content guided by modification indices. The standardized path diagram of the modified model is presented in Fig. 1. All fit indicators for the modified model achieved acceptable value, as detailed in Table 4.

Table 4
Measures of the initial and modified models

Fit indicators	Initial model	Modified model	Acceptable value
χ^2/df	2.526	2.171	< 3
CFI	0.893	0.919	≥ 0.90
TLI	0.876	0.905	≥ 0.90
RMSEA	0.072	0.063	≤ 0.08

Note: *CFA* Confirmatory factor analysis; χ^2/df Residual degrees of freedom; *CFI* Comparative fit index; *TLI* Tucker-Lewis index; *RMSEA* Root mean square error of approximation.

A multi-group confirmatory factor analysis was conducted to evaluate measurement invariance across gender and education level groups. Configural, weak, and strong invariance were tested sequentially. Model comparisons were based on changes in fit indices ΔCFI and ΔRMSEA, with invariance supported if |ΔCFI| ≤ 0.01 and |ΔRMSEA| ≤ 0.015 [22]. As shown in Table 5, both grouping variables met these criteria, indicating scalar invariance and supporting valid cross-group comparisons among medical students.

Table 5
Measurement invariance across gender and education level

	CFI	RMSEA	|ΔCFI|	|ΔRMSEA|
		Gender		
Structural	0.891	0.053		
Weak	0.890	0.052	0.001	0.001
Strong	0.889	0.051	0.001	0.001
		Education level		
Structural	0.905	0.050		
Weak	0.903	0.049	0.002	0.001
Strong	0.898	0.049	0.005	0.000

3.5. Cronbach's alpha reliability

To assess the internal consistency reliability of the scale, Cronbach's alpha coefficients were calculated (Table 6). The overall Cronbach's alpha coefficient for the scale was 0.902, and the coefficients for the individual dimensions ranged from 0.787 to 0.840. All values exceeded the threshold of 0.7, indicating that the scale possesses good internal consistency. Additionally, the descriptive statistics for each factor are summarized in Table 6.

Table 6
Descriptive statistics and reliability of factors

Factors	Mean	S.D.	Skewness	Kurtosis	Cronbach *α*
1. Technology and application	21.98	5.983	-0.170	-0.267	0.832
2. Humanistic values	31.59	5.348	-0.596	0.906	0.787
3. Ethics and fairness	25.91	5.113	-0.399	-0.101	0.811
4. AI-enhanced learning	21.28	4.149	-0.821	1.125	0.840
Total	100.76	16.166	-0.39	0.192	0.902

Discussion

Grounded in the systematic requirements for the core competencies of medical students in the AI era, this study developed the AILS-MS through a rigorous process. The resulting scale demonstrated excellent reliability (Cronbach's α = 0.902), and CFA indicated a satisfactory model fit (χ^2/df = 2.171; CFI = 0.919; RMSEA = 0.063). The results indicate that the scale exhibits robust psychometric properties in terms of construct validity, internal consistency, and item discrimination, establishing a solid measurement foundation for the scientific assessment of AI literacy among medical students.

In contrast to existing general AI literacy scales that predominantly focus on technical dimensions such as tool utilization and computational thinking [7, 18], the present scale is based on *the UNESCO framework* and deeply integrates the professional characteristics and ethical imperatives of medical education. It embodies a four-dimensional structure that organically unifies technological accessibility, the irreplaceability of humanistic values, ethical prudence, and learning autonomy. Within the technology and application dimension, items such as "using effective prompts to guide AI systems toward generating desired outputs" and "leveraging AI to devise innovative solutions when facing complex clinical or academic challenges" transcend a simplistic assessment of AI tool functionality. Instead, these items emphasize AI as an intermediary bridging clinical practice, scientific research, and humanistic care. Furthermore, this study establishes humanistic values and ethics as distinct primary indicators, emphasizing that medical students should possess humanistic spirit and critical thinking capabilities regarding issues like algorithmic bias during AI-assisted diagnosis and treatment. This approach not only upholds medical professionalism but also represents its necessary extension in the era of intelligence.

The novel dimension "AI-enhanced learning" proactively addresses the rapid iteration of medical knowledge and the evolution of diagnosis and treatment models within the context of the global digital transformation of medical education[23]. This dimension emphasizes the urgent need for medical students to move beyond the passive use of AI and actively integrate it into their personal learning ecosystems and professional development. Through items such as "applying AI in interdisciplinary settings" and "viewing AI as a vital tool supporting my commitment to lifelong learning," the scale

precisely captures the ability to apply AI in medical settings to foster interdisciplinary collaboration, expand cognitive boundaries, and sustain continuous growth. This not only aligns with the future trend of AI-driven interprofessional clinical learning but also highlights the empowering role of AI in fostering autonomous learning, integrating diverse knowledge, and supporting lifelong development, thereby meeting the core requirements for cultivating medical talent in the era of intelligence [24].

The framework of AI literacy for medical students established by this scale represents a systematic integration of the core requirements emphasized in the medical profession, including humanistic care, ethical prudence, and interdisciplinary collaboration. High-level AI literacy extends beyond mere technical proficiency. It entails the capacity to utilize AI creatively, examine it critically, and manage it responsibly within complex clinical scenarios. To effectively enhance the AI literacy of medical students, we recommend prioritizing the following three aspects in medical education and policy formulation: First, strengthen the cultivation of practical competencies. Medical curricula should systematically incorporate AI application training based on simulated cases to improve the ability of students to reasonably deploy, verify, and integrate AI outputs in authentic diagnostic and therapeutic environments. Second, training in critical thinking requires reinforcement. Educators should utilize the teaching of “AI error cases” to guide students in identifying hallucinations, logical inconsistencies, and algorithmic biases, thereby fostering independent judgment regarding AI-generated results. Third, emphasize education concerning ethics and law. Integrate AI-related laws and regulations into mandatory coursework to clarify the compliance boundaries for the use of medical AI and to reinforce awareness of data security and legal liability.

Conclusion

Guided by *the UNESCO framework* and incorporating the new requirements for talent cultivation in the AI era medical education, this study developed and validated an AI literacy assessment tool tailored for medical students. The methodology employed literature review, semi-structured interviews, and two rounds of Delphi expert consultations. The AILS-MS consists of 4 primary indicators and 20 items, demonstrating robust reliability, validity, and structural stability. This scale serves as an effective instrument for the scientific assessment of medical students’ AI literacy. Furthermore, it provides a foundation for curriculum design, educational reform, and competency development research, thereby promoting the deep integration of AI into medical education and its systematic advancement.

As an AI literacy measurement tool specifically focused on medical scenarios, this scale has diverse applications. At the individual level, the scale enables students to accurately evaluate their own AI literacy and identify gaps in knowledge and skills. At the instructional level, it serves as an effective instrument for medical schools to conduct baseline surveys and track the effectiveness of AI literacy education, providing empirical support for curriculum optimization and teaching reform. In terms of exploring psychological mechanisms, the scale can be utilized alongside variables such as AI usage frequency, AI anxiety levels, and willingness to pay to systematically examine the pathways influencing AI literacy. Additionally, it allows for the analysis of associations between AI literacy and mental health

indicators like social phobia and difficulties in emotion regulation, providing a scientific basis for early intervention.

Declarations

Acknowledgements

We are grateful to all those who took part in or assisted with this study project.

Authors' contributions

Weijia Feng: Conceptualization, data curation, software, validation, formal analysis, visualization, writing-original draft and writing-review & editing. Yaoyue Hu: Conceptualization, methodology, writing-original draft, writing-review & editing and funding acquisition. Jialei Tang and Han Zhou: Conceptualization, data curation, writing-review & editing and supervision. Hong Xu: Conceptualization, methodology, supervision, project administration and funding acquisition.

Funding statement

This research was supported by the Chongqing Medical University Medical Education Research Planning Project (Grant No. J0625003); and Chongqing Higher Education Scientific Research Project (Grant No. cqgj25024C).

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics approval and consent to participate

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki. It involved a retrospective analysis of data collected from participants who had provided informed consent. All data were anonymized during collection and analysis to protect participant privacy. The study received approval from the Ethics Committee of Chongqing Medical University (No.2026-145).

Consent for publication

Not applicable.

Conflicts of interest

The authors declare no competing interests.

References

1. Morley J, DeVito NJ, Zhang J. Generative AI for medical research. BMJ. 2023;382:1551.
2. Bates DW, Levine D, Syrowatka A, Kuznetsova M, Craig KJT, Rui A, et al. The potential of artificial intelligence to improve patient safety: a scoping review. npj Digital Medicine. 2021;4:54.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25:44–56.
4. Park SH, Do K-H, Kim S, Park JH, Lim Y-S. What should medical students know about artificial intelligence in medicine? J Educ Eval Health Prof. 2019;16:1149130.
5. United Nations Educational, Scientific and Cultural Organization (UNESCO). AI competency framework for students. Paris: UNESCO Publishing; 2024.
6. Gong M, Pan H, Liu H, Tong K, Li J, Ma Y, et al. Expert Consensus on the Artificial Intelligence Proficiency Competency List and Assessment Framework for Medical Students (2025 Edition). Acta Academiae Medicinae Sinicae. 2026;:1–11.
7. Zhao Y, Peng X, Cun G. Artificial Intelligence Literacy: Connotation, Framework, and Evaluation System. Library and Information Service. 2025;69:83–93.
8. Zhou Q, Xu Y, Cai Y. An Examination of AI Literacy Among University Students: Current Status and Multifaceted Influencing Factors. Documentation, Information & Knowledge. 2024;41:38–48.
9. Ma S, Chen Z. The Development and Validation of the Artificial Intelligence Literacy Scale for Chinese College Students (AILS-CCS). IEEE Access. 2024;12:146419–29.
10. Ding J, Guo L, Zhang W, Liu H. Development of an Artificial Intelligence Literacy Indicator System for University Students Using a Combined AHP-BPNN Approach. Journal of Distance Education. 2025;43:46–56.
11. Li N, Liu S, Ji J. Construction of AI Literacy Evaluation System for College Students from the Continuum Perspective. Information Studies: Theory & Application. 2025;48:150–9.
12. Laupichler MC, Aster A, Meyerheim M, Raupach T, Mergen M. Medical students' AI literacy and attitudes towards AI: a cross-sectional two-center study using pre-validated assessment instruments. BMC Med Educ. 2024;24:401.
13. Karaca O, Çalışkan SA, Demir K. Medical artificial intelligence readiness scale for medical students (MAIRS-MS) – development, validity and reliability study. BMC Med Educ. 2021;21:112.
14. Ma S, Chen Z. The Development and Validation of the Artificial Intelligence Literacy Scale for Chinese College Students (AILS-CCS). IEEE Access. 2024;12:146419–29.
15. Han S, Zhang Y. Developing a validated assessment of artificial intelligence literacy for Chinese university students based on educational objectives taxonomy. Journal of Research on Technology in Education. 2025;0:1–16.
16. Wu D, Sun X, Liang S, Qiu C, Wei Z. Construction of AI Literacy Evaluation System for College Students and an Empirical Study at Wuhan University. Front Digit Educ. 2025;2:6.
17. Liu M, Guo S, Su Y, Li Y, Kui X. Development and application of the assessment tool for college students' generative artificial intelligence literacy. The Modern Distance Education Research.

2025;37:24-33+70.
18. Lee SC, Baby T, Vongvit R, Lee J, Kim YW, Cha MC, et al. Development and validation of Generative AI Competence Scale (GenAIComp) among university students. Technology in Society. 2025;84:103059.
19. Skulmoski GJ, Hartman FT, Krahn J. The Delphi method for graduate research. Journal of Information Technology Education: Research. 2007;6:1–21.
20. He H, Zhou T, Zeng D, Ma Y. Development of the competency assessment scale for clinical nursing teachers: Results of a Delphi study and validation. Nurse Education Today. 2021;101:104876.
21. Marsh HW, Morin AJS, Parker PD, Kaur G. Exploratory Structural Equation Modeling: An Integration of the Best Features of Exploratory and Confirmatory Factor Analysis. Annu Rev Clin Psychol. 2014;10:85–110.
22. Cheung GW, Rensvold RB. Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. Structural Equation Modeling: A Multidisciplinary Journal. 2002;9:233–55. https://doi.org/10.1207/S15328007SEM0902_5.
23. WHO. World health statistics 2023: monitoring health for the SDGs, sustainable development goals. World Health Organization; 2023.
24. Shaw K, Henning MA, Webster CS. Artificial Intelligence in Medical Education: a Scoping Review of the Evidence for Efficacy and Future Directions. MedSciEduc. 2025;35:1803–16.

Figures

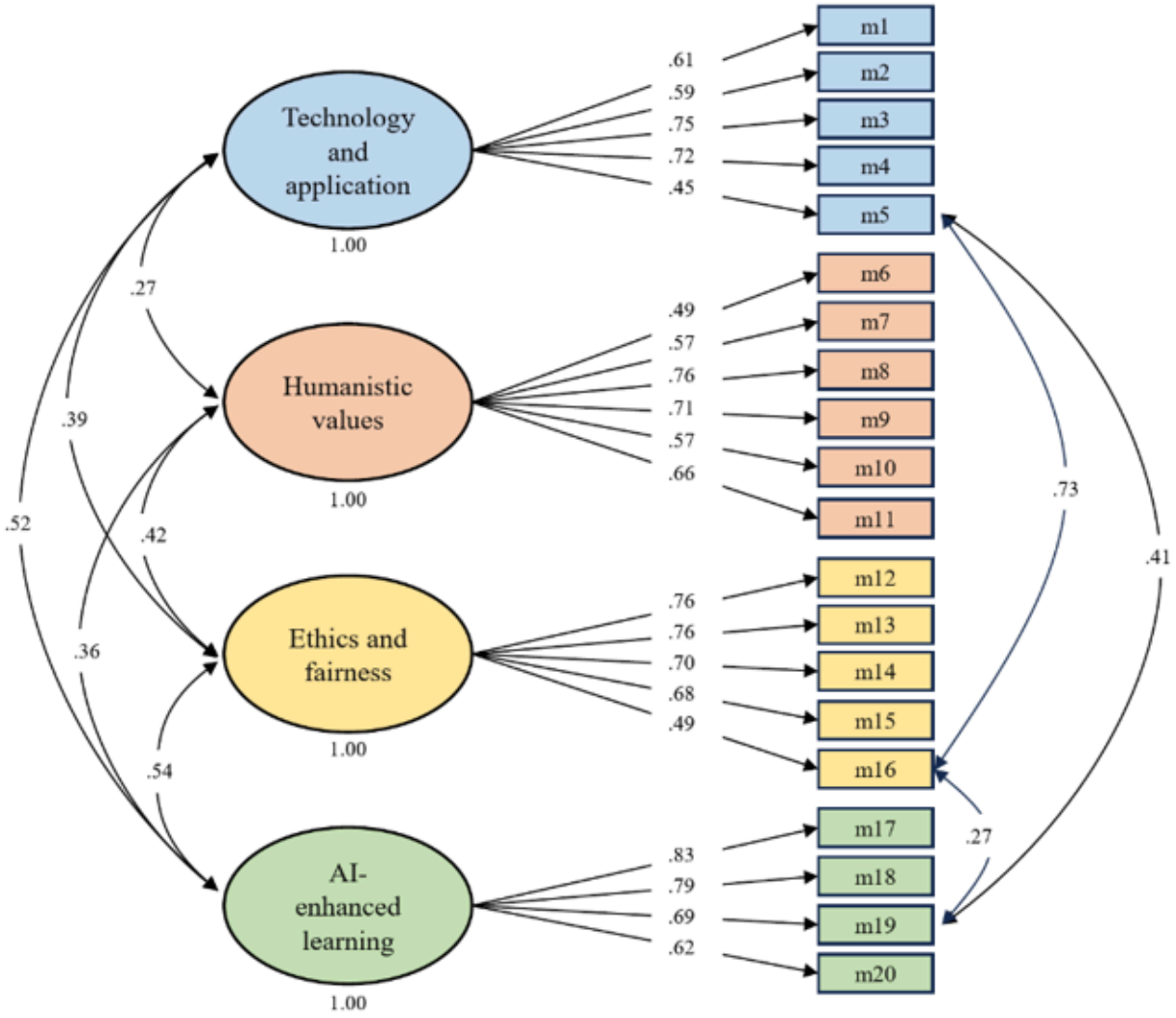


Figure 1

Path diagram about confirmatory factor analysis

Supplementary Files

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

- Supplementaryfile1.docx
- Supplementaryfile2.docx