

医学生生成式人工智能素养教育研究

徐铭超, 谢礼丹, 谢启龙

(眉山市中医医院, 四川眉山 620000)

[摘要]目的:梳理生成式人工智能背景下医学生人工智能素养教育研究进展,分析其在医学教学评价中的应用及启示。方法:参照范围综述思路,检索中文数据库、英文数据库及人工智能治理文件,纳入相关文献。结果:共纳入59篇文献,生成式人工智能已用于命题、病例分析、问诊训练、作业评阅、病历书写评价等场景。结论:教学评价能够集中呈现医学生使用生成式人工智能的知识、技能与态度要求,可作为课程建设、教学实施、测评改进和治理规范的切入点。

[关键词]生成式人工智能;医学生;人工智能素养;医学教育;教学评价

[作者简介]徐铭超(1995—),男,山西临汾人,眉山市中医医院主治中医师,医学硕士,研究方向:中医眼科学。谢礼丹(1987—),女,四川眉山人,眉山市中医医院副主任中医师,医学硕士,研究方向:中医眼科学。谢启龙(1984—),男,四川眉山人,眉山市中医医院副主任医师,医学硕士,研究方向:眼科学。

[DOI] <https://doi.org/10.62662/kxwxz0307026>

[中图分类号] G642

[本刊网址] www.oacj.net

[投稿邮箱] jkw1966@163.com

引言

生成式人工智能正在改变医学生获取知识、完成作业和参与课堂训练的方式。与既往主要用于影像识别、数据建模或临床决策支持的人工智能不同,大语言模型能够生成病例材料、模拟问诊对话、解释医学概念、辅助命题、评价文本并提供学习反馈,因而更容易进入医学教育的日常环节。近年国内医学教育研究中,DeepSeek 辅助命题、医学生反思性写作评阅、住院医师病例分析考核、眼科大病历书写评价、医学检验思维能力评价、诊断学问诊教学和虚拟标准化病人等文献逐渐增多。这说明生成式人工智能对医学教育的影响不仅限于提高学习效率,也体现在教学评价方式的变化。

医学生人工智能素养既包括对人工智能基本原理、适用边界和伦理规范的理解,也包括在医学情境中使用、评价和约束人工智能输出的能力。现有研究多从课程设置、能力框架或使用态度方面来讨论人工智能素养教育,但对“如何判断医学生是否真正具备生成式人工智能素养”的讨论仍不充分。学生面对人工智能生成的诊断建议、病例分析或患者沟通文本时,不能只判断表达是否流畅,还要判断证据是否充分、推理是否完整、隐私是否受

保护、结论是否适合进入学习或临床训练情境。因此,本文以教学评价重构为切入点,梳理生成式人工智能背景下医学生人工智能素养教育的研究现状,并在课程建设、教学方法、测评体系和伦理治理方面讨论其带来的影响和启发。

一、资料与方法

(一) 确定研究问题

围绕生成式人工智能进入医学教育后的教学评价变化,本文拟回答以下问题:(1)生成式人工智能背景下医学生人工智能素养的核心内涵是什么;(2)生成式人工智能主要进入了哪些医学教学评价场景;(3)课程建设、教学方法、测评体系和伦理治理如何支撑人工智能素养培养;(4)基于教学评价重构,医学生人工智能素养教育可形成哪些启发。

(二) 检索策略

中文数据库检索中国知网、万方和维普,英文及开放来源检索 PubMed、Europe PMC、OpenAlex、Crossref,并补充 WHO、UNESCO、AAMC 等机构发布的人工智能治理和医学教育相关文件。中文检索词包括“医学生”“医学教育”“人工智能素养”“生成式人工智能”“大语言模型”“ChatGPT”“DeepSeek”“教学评价”“测评体系”“课程建设”

“伦理治理”等;英文检索词包括“medical students”“artificial intelligence literacy”“generative AI”“large language models”“undergraduate medical education”“curriculum”“assessment”“ethics”“governance”等。检索时采用主题词与自由词结合,并根据不同数据库语法进行调整。检索时限为各数据库建库至2026年6月5日。

(三)文献纳入与排除标准

参照PCC原则确定文献范围。纳入标准:(1)研究对象为医学生、住院医师、医学教师或医学教育相关群体;(2)主要概念涉及人工智能素养、生成式人工智能素养、生成式人工智能医学教育、教学评价、测评工具或伦理治理;(3)使用情境为医学课程、临床前训练、住院医师培训、医学教学测评、医学人文或科研训练。排除标准:(1)纯临床模型性能、患者服务或医疗管理研究,且不能支撑医学教育问题;(2)普通高校人工智能课程或通用数字素养研究,缺少医学情境;(3)重复题录、勘误、预印本或题录信息不完整文献;(4)护理、助产等相邻专业文献仅作为背景参照,不纳入核心分析。

(四)文献筛选与资料提取

将检索记录统一整理后,依据题名、作者、DOI和来源信息去重。由研究者依据纳入与排除标准阅读文题和摘要进行初筛,再阅读全文或可获得的核心内容进行复筛,并由研究团队对筛选结果进行复核;意见不一致时通过讨论解决。对于政策文件、专家共识和机构指南,重点依据发布机构权威性、内容与医学教育或人工智能治理的相关性进行判断。资料提取内容包括作者、年份、来源、研究对象、研究类型、生成式人工智能应用场景、人工智能素养要素、教学或测评方式、伦理治理要求及主要结论。主题归纳时重点关注能力框架与测评、教学评价场景、课程与教学设计、伦理治理和培养启示。

二、结果

(一)文献筛选结果

初步获得记录735条,其中中文数据库145条,包括中国知网53条、万方71条、维普21条;英文及开放来源590条。剔除题录核心信息缺失记录12条后,进入去重记录723条;去重后剩余661条。阅读题名和摘要排除189条,472条进入复筛;复筛时根据研究对象、医学教育情境、生成式人工智能相关性及全文或核心内容可获得性继续排除413条,最终纳入59篇文献。文献筛选流程见图1。

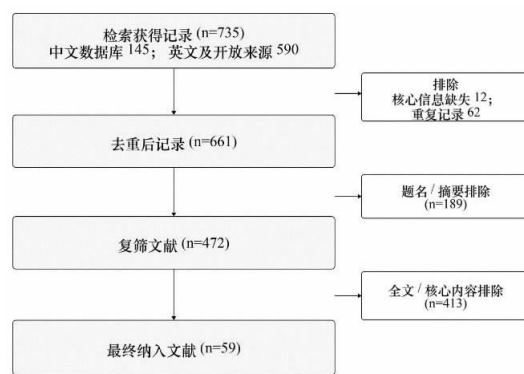


图1 文献筛选流程图

(二)纳入文献的基本特征

59篇纳入文献发表于2021—2026年,其中2025—2026年文献占多数,提示该主题近期发展较快。从来源看,中文文献38篇,英文及开放来源文献21篇。从文献类型看,教学实践或评价研究25篇,理论观点或路径研究14篇,共识、指南或政策文件7篇,综述或证据整合5篇,调查、量表或框架研究4篇。主题方面,能力框架与测评22篇,生成式人工智能参与医学教学评价18篇,医学课程与教学实践11篇,伦理治理与责任边界5篇,医学教育应用与影响2篇,通用教育治理参考1篇。

(三)医学生人工智能素养的内涵

纳入文献普遍提出,医学生人工智能素养不宜简化为是否会使用某一工具,而应该在能否使用得好方面也提出要求。我国专家共识将医学生人工智能素养划分为知识、技能和态度维度,并提出能力清单与测评框架。医学教师人工智能素养专家共识和教师人工智能素养测评量表研究也表明,教师的课程设计、评价引导和伦理把关能力,是学生人工智能素养培养的重要前提。国外文献同样强调,医学生需要理解人工智能基本机制、适用边界和偏倚来源,并在医学情境中进行批判性使用。

生成式人工智能使这一素养内涵进一步具体化,学生不仅向工具提问,还会接收模型生成的疾病解释、病例分析、学习计划、科研文本、命题材料和评价反馈。相应地,素养教育需要从“会提问”延伸到“会核查、会修订、会说明”。应该包含知识、态度、技能等多维度要求,知识维度应包括模型局限、医学证据层级、数据安全和基本伦理规范;技能维度应包括提示设计、证据检索、输出审查、错误识别和人机协同表达;态度维度应强调审慎信任、患者安全、学术诚信和职业责任。因此,与通用人工智

能素养相比,医学生人工智能素养的专业特征在于,必须把模型输出放回医学证据、教学目标和临床责任框架中加以判断。

(四)生成式人工智能参与医学教学评价的主要场景

纳入文献中,教学评价相关研究数量较多,且应用场景较为具体。国内研究已涉及 DeepSeek 辅助前列腺癌教学测评命题、大语言模型评阅医学生反思性写作作业、住院医师理论及病例分析考核、重症医学住院医师规范化培训考核、眼科专科大病病历书写评价、医学检验本科生思维能力教学评价和医学检验教学评价等。相关研究显示,生成式人工智能能够提高材料生成、初步反馈和评分辅助的效率,但其输出质量受任务描述、评分标准、模型版本、医学语境和教师复核机制影响,尚不能作为独立评价主体。

重点关注教学评价场景,是因为它能够较集中地呈现学生人工智能素养的实际水平。学生在病例分析、文献阅读、作业修改或病历书写中使用生成式人工智能时,可能将语言流畅误认为医学正确,将模型建议误认为证据结论,也可能忽视患者隐私和学术诚信要求。若评价只关注最终答案是否完整,而不关注学生如何使用、核查、修订和说明人工智能输出,就难以判断其是否具备医学情境下的人工智能素养。因此,医学教学评价可逐步纳入人工智能输出审查、证据链标注、模型建议采纳说明、错误输出修订和使用声明等任务。

(五)课程与教学

现有课程研究显示,生成式人工智能可用于医学影像学、人体解剖学、诊断学、医学大数据分析、留学生医学课程和临床医学本科教学等场景。但如果课程建设主要停留在工具介绍、课堂互动或学习材料生成,学生获得的往往是零散使用经验,难以形成稳定的评价判断能力。综合来看,较为适宜的路径是以评价任务反推课程设计,将人工智能素养嵌入医学信息学、循证医学、诊断学、临床技能训练、科研训练和医学伦理课程。

具体而言,低年级课程可重点讲授生成式人工智能基本原理、模型局限、提示设计和数据边界;临床前课程可围绕病例分析、问诊脚本、影像描述、药物信息和文献证据训练输出核查;临床实践或住院医师培训阶段可在教师监督下开展病历书写评价、病例讨论、患者沟通材料修订和客观结构化临床考

试(objective structured clinical examination, OSCE)情境训练。教学方法上,可采用模型回答比对、错误输出纠正、证据链追踪、反思日志和人机协同病例讨论等形式。这样设计的重点不是增加一门孤立的人工智能课程,而是让学生在不同学习阶段反复经历“提出问题—生成输出—核查证据—修订说明—伦理反思”的过程。

(六)测评体系与伦理治理

生成式人工智能参与医学教育后,传统以结果为中心的评分方式面临新的挑战。学生提交的答案可能经过模型辅助,教师也可能使用模型生成反馈或试题。如果评价制度没有明确人工智能使用边界,学生的真实能力、学术诚信和患者安全意识都难以判断。WHO 关于健康领域大模型治理的指南、AAMC 关于医学课程负责任使用人工智能的原则,以及 UNESCO 关于生成式人工智能教育研究的指导文件,均强调透明、隐私、公平、安全、人类监督和问责。国内关于生成式医学人工智能临床应用伦理治理、医疗决策伦理困境和中国实践路径的研究,也提示医学教育不能将治理问题推迟到临床阶段。

因此,医学院校应将伦理治理转化为可执行的评价规则。评价标准中应明确哪些病例、影像、病历和科研数据不得输入开放模型,哪些作业允许使用人工智能辅助,使用后如何声明,模型生成的反馈、命题和评分结果如何复核,学生如何说明人工智能输出的采纳依据。测评方式也应从单纯评价答案质量,扩展为评价学生是否完成证据核查、是否识别模型错误、是否能够说明修订理由、是否遵守隐私和学术规范。只有将治理要求写入课程大纲、作业说明和评价量规,学生才可能在明确规则下形成负责任的使用习惯。

三、讨论

生成式人工智能背景下医学生人工智能素养教育的切入点,正在从工具介绍转向教学评价场景。与通用人工智能素养相比,医学教育更关注模型输出是否符合医学证据、是否保护患者权益、是否有助于临床思维形成,以及使用者能否说明责任边界。教学评价环节要求学生呈现思考过程和判断依据,因而能够较直接地观察其人工智能素养水平。

当前研究仍存在若干不足。第一,课程建设研究较多,但不少实践仍以工具使用和课堂效率为中心,尚未充分说明学生通过课程形成了哪些可评价

能力。其原因在于生成式人工智能工具迭代速度较快,单次课程实践容易围绕新工具功能展开,而难以沉淀为稳定的课程目标和评价指标。第二,测评研究逐渐增多,但部分研究主要关注模型能否完成命题、评分或反馈,对学生如何学习评价人工智能输出讨论不足。医学教育评价同时涉及医学知识、临床推理、沟通表达、伦理判断和职业态度,通用人工智能素养量表难以直接迁移到医学场景。第三,真实临床情境中的研究仍相对有限,患者隐私、数据安全、带教责任和学术诚信规则尚未充分进入教学评价设计。其背后既有技术和伦理风险,也与院校层面的人工智能使用规范、数据授权机制和教师培训不足有关。第四,本土化测评工具仍需完善。已有医学生人工智能素养能力清单为评价提供了基础,但如何将其转化为课程量规、OSCE任务、病例分析评分表和纵向追踪指标,仍需要进一步研究。

因此,后续教改研究可将“评价人工智能输出的能力”作为连接课程、教学、测评和治理的指标之一。研究重点不应只是比较人工智能与教师、学生谁的得分更高,而应进一步考查学生能否在人工智能参与的任务中表现出证据核查、错误识别、解释说明、伦理判断和责任意识。这样的评价设计更符合医学教育强调过程性、情境性和职业性的特点。

四、基于教学评价重构的培养启示

基于教学评价重构的培养启示,并不是另行设计一套脱离课程的培养方案,而是结合纳入文献中的课程、测评和治理建议,将人工智能素养要求嵌入医学生已有培养进程中。结合医学教育阶段特点,可从基础认知、证据核查、人机协同评价和临床责任判断四个层面归纳培养启示。

这一培养思路的重点在于受控使用和可解释评价。低年级阶段应避免学生把生成式人工智能视为直接给出答案的工具,重点建立规则意识和边界意识;临床前阶段应通过病例和文献任务训练证据核查能力;进入临床训练后,应在教师监督下处理病历、问诊、患者沟通等更接近真实医疗服务的任务,强调隐私保护和责任说明。对于有科研或创新训练需求的学生,还可引导其参与模型评价、课程改进和数据治理,使人工智能素养从一般使用能力进一步发展为医学教育创新能力。

五、小结

生成式人工智能正在影响医学教育的知识获

取方式,也在改变教学评价内容。医学生人工智能素养教育若停留在工具演示和提示词训练方面,就难以回应医学教育对证据、患者安全和职业责任的要求。以教学评价重构为切入,可以把课程建设、教学方法、测评体系和伦理治理连接起来,使学生在病例分析、问诊训练、作业评阅、命题辅助和病历评价等任务中形成可观察、可评价的人工智能素养。后续研究应重点开展本土化测评工具开发和真实教学场景中的纵向评价研究,为新医科背景下复合型医学人才培养提供更具体的证据支持。

参考文献:

- [1]刘栩豪,陈敏丰,蔡焱. DeepSeek 辅助命题用于前列腺癌教学测评的可行性研究[J]. 中华医学教育探索杂志, 2026, 25(2): 137-143.
- [2]柴桦,闫昱江,卿平. 从工具替代到人机协作:深度推理大语言模型评阅医学生反思性写作作业的个案研究[J]. 中华医学教育探索杂志, 2026, 25(1): 1-11.
- [3]何源龙,毛鸣. 基于 LLM 模型的虚拟标准化病人在骨科住院医师培训中的应用效果评估[J]. 泰州职业技术学院学报, 2026, 26(1): 16-18, 26.
- [4]王晨,齐惠颖. 基于生成式人工智能的“医学大数据分析”课程创新教学模式研究与实践探索[J]. 中国医学教育技术, 2026, 40(3): 358-365.
- [5]周元凯,刘培,刘圣均,等. 生成式人工智能在重症医学住院医师规范化培训考核中的应用探索[J]. 协和医学杂志, 2026, 17(1): 286-291.
- [6]夏亮,于雯雯,邹多宏,等. 生成式人工智能赋能转化医学导向的口腔颌面外科教学价值、挑战与策略初探[J]. 中国实用口腔杂志, 2026, 19(2): 218-223.
- [7]梁晨,李瑞春,荆江鹏,等. ChatGPT 辅助 BOPPS 教学方法在留学生神经外科见习教学中的应用[J]. 中国医学教育技术, 2025, 39(3): 379-385.
- [8]容晓莹,徐懋,周阳,等. ChatGPT 辅助基于案例学习教学在八年制临床医学专业学生麻醉学教学中的应用[J]. 中华医学教育杂志, 2025, 45(8): 610-615.
- [9]张露,丁海,徐学静. 人工智能技术驱动的医学检验教学评价[J]. 标记免疫分析与临床, 2025, 32(12): 2478-2483.
- [10]袁志怡,裴志怡,林佳艺,等. 医学生对生成式人工智能融入医学教育的认知——一项 Meta 整合研究[J]. 医学与哲学, 2025, 46(14): 52-57.
- [11]杨崇清,龙燕,崔红光,等. 基于生成式人工智能技术的眼科专科大病历书写评价研究——以 DeepSeek 模型为例[J]. 中国高等医学教育, 2025(6): 1-3.
- [12]于运亮,矫秀,李婷,等. 生成式 AI 赋能医学检验本科思维力教学评价的路径探索[J]. 标记免疫分析与临

床,2025,32(12):2471-2477,2605.

[13]朗柳玲,苏金婷,任文宁,等.生成式AI赋能高等医学教育:从课程到评价的变革之路[J].中国医学教育技术,2025,39(6):698-704.

[14]辛晓明,宋晓蕾,韩秋琴,等.生成式人工智能促进医学教学改革的路径探索——基于医学院大学生的调研分析[J].卫生职业教育,2025,43(22):105-108.

[15]周元凯,孙俊,刘圣均,等.生成式人工智能在住院医师理论及病例分析考核中的应用与挑战[J].协和医学杂志,2025,16(5):1352-1356.

[16]任晓阳,朱姣,刘亚萍,等.生成式人工智能在诊断学问诊教学中的应用[J].中国医学教育技术,2025,39(1):29-32.

[17]陆璐,丁峰,肖顺华.生成式人工智能技术对医学教育教学评价的影响及应对建议思考[J].湖北开放职业学院学报,2025,38(10):158-160.

[18]魏兵,林小慧,陈素一,等.生成式人工智能赋能地方医学院校临床医学博士培养体系构建[J].医学教育研究与实践,2025,33(6):761-766.

[19]丛尚,白沃涵,陈钟,等.大语言模型在病例生成、个性化教学与评估中的研究进展[J].医学分子生物学杂志,2026,23(1):97-106.

[20]弓孟春,潘慧,刘辉,等.医学生人工智能素养能力清单与测评框架专家共识(2025年版)[J].中国医学科学院学报,2026,48(1):13-23.

[21]郑宇,董芷菁.人工智能背景下医科类专业人才核心素养提升路径[J].中国就业,2026(5):92-93.

[22]陈第华.人工智能赋能医学生人文素养培育的限度与出路[J].南京医科大学学报(社会科学版),2026,26(3):254-260.

[23]潘慧,弓孟春,卢景辉,等.医学教师的人工智能素养专家共识(2025年版)[J].中国医学科学院学报,2026,48(1):1-12.

[24]王娜,龚玉雯,宋飞雪.数智化教学与人工智能在临床医学教育中的应用与发展[J].继续医学教育,2026,40(2):26-31.

[25]申正付,熊韵波,杨秀木,等.高等医学院校教师AI素养测评量表的编制及应用研究[J].蚌埠医科大学学报,2026,51(3):306-314.

[26]张思玮.人工智能时代,如何培育复合型医学人才?《医学生AI素养能力清单及其测评框架专家共识(2025)》发布[EB/OL]. <https://news.sciencenet.cn/htmlnews/2025/8/550377.shtm>,2025-8-27.

[27]陈峰.人机协同医学教育模式研究[D].武汉:华中师范大学,2025.

[28]宋雨欣,李子怡,丁宁.生成式人工智能对医学教育的影响探析[J].中国医学教育技术,2024,38(3):281-286.

[29]Li C., Tsuei S. H. T., Wu H. AI Literacy among

Chinese Medical Students: Cross-sectional Examination of Individual and Environmental Factors [J]. *JMIR Medical Education*, 2026, 12.

[30]Catan-Inan F. Artificial Intelligence Literacy and Readiness in Future Health Care Professionals: A Cross-sectional Study [J]. *Croatian Medical Journal*, 2026, 67(1): 4-12.

[31]Gong M. C., Pan H., Liu H., et al. Expert Consensus on the Artificial Intelligence Proficiency Competency List and Assessment Framework for Medical Students (2025 Edition) [J]. *Acta Academiae Medicinae Sinicae*, 2026, 48(1): 13-23.

[32]Chang L., Sreedhar R. Integrating AI Literacy into Medical Education: Preparing Future Clinicians for an AI-driven Healthcare System [J]. *Medical Science Educator*, 2026, 36(1): 17-22.

[33]Yousef M., Deeb S., Alhashlamon K. AI Usage among Medical Students in Palestine: A Cross-sectional Study and Demonstration of AI-assisted Research Workflows [J]. *BMC Medical Education*, 2025, 25(1): 693.

[34]Simoni J., Urtubia-Fernandez J., Mengual E., et al. Artificial Intelligence in Undergraduate Medical Education: An Updated Scoping Review [J]. *BMC Medical Education*, 2025, 25(1): 1609.

[35]Huang M. Empowering Medical Students with AI Literacy: A Curriculum Development Journey [J]. *Medical Education*, 2025, 59(5): 550-551.

[36]Haider M. H., Majeed M. U., Irfan M. Needs of Digital Literacy in the Era of Artificial Intelligence: A Comprehensive Review [J]. *Inverge Journal of Social Sciences*, 2025, 4(4): 396-409.

[37]Ahmed W. M., Abdalmotalib M. M., Elbadawi M. H., et al. Shaping the Future of Medical Education: A Cross-sectional Study on ChatGPT Attitude and Usage among Medical Students in Sudan [J]. *PLoS ONE*, 2025, 20(5).

[38]Ajalo E., Mukunya D., Nantale R., et al. Widespread Use of ChatGPT and Other Artificial Intelligence Tools among Medical Students in Uganda: A Cross-sectional Study [J]. *PLoS ONE*, 2025, 20(1).

[39]Roberts L. W. Addressing the Novel Implications of Generative AI for Academic Publishing, Education, and Research [J]. *Academic Medicine*, 2024, 99(5): 471-473.

[40]Sridharan K., Sequeira R. P. Artificial Intelligence and Medical Education: Application in Classroom Instruction and Student Assessment Using a Pharmacology & Therapeutics Case Study [J]. *BMC Medical Education*, 2024, 24(1): 431.

[41]李琨,贾立伟,石晓明.“教育信息化2.0”背景下医学高专院校《人工智能通识课》课程建设方案初探[J].甘肃

科技,2021,37(18):80-82,35.

[42]洪敏萍,黎良山,国海燕,等. ChatGPT 辅助医学影像学教学的探索与实践[J]. 中国高等医学教育,2026(1):108-109,118.

[43]陈珊珊,毕昕,李伟,等. 数智赋能背景下的医学影像设备学课程建设路径探索与实践[J]. 中国医学教育技术,2026,40(2):228-233.

[44]朱惠敏,胡蝶,郭振宇,等. 生成式人工智能在人体解剖学教学实践中的应用[J]. 基础医学教育,2026,28(2):167-170.

[45]李子广,唐佳鑫,周佳慧,等. ChatGPT 联合 ASSURE 评估在临床医学本科教学中的探索与实践[J]. 护理学,2025,14(12):2456-2462.

[46]Taylor A. W. R. , Patel F. Y. , Smyth E. K. , et al. AI in the Classroom: Observing Preclinical Students' Use of ChatGPT during Case-based Learning at a UK Medical School [J]. *The Clinical Teacher*, 2026, 23(2).

[47]Li J. , Ai F. , Wang J. , et al. Application of AI-generated Content in Medical Education: Systematic Review of the Impact on Critical Thinking Abilities of Medical Students[J]. *JMIR Medical Education*, 2026, 12.

[48]Feigerlova E. How to Teach Generative Artificial Intelligence in Undergraduate Medical Education [J]. *The Clinical Teacher*, 2026, 23(3).

[49]Ma X. , Liu X. , Sun H. The Effectiveness of GAI-assisted Teaching Methods in Medical Education: A Systematic Review and Meta-analysis [J]. *Frontiers in Public Health*, 2026, 14.

[50]Kim D. H. , Kang Y. J. , Lee Y. M. Twelve Tips for Developing and Implementing AI Curriculum for Undergraduate Medical Education[J]. *Medical Education Online*, 2025, 30(1).

[51]Hale J. , Alexander S. , Wright S. T. , et al. Generative AI in Undergraduate Medical Education: A Rapid Review [J]. *Journal of Medical Education and Curricular Development*, 2024, 11.

[52]Association of American Medical Colleges. Principles of Responsible AI for Medical School and Residency Curricula [Z]. 2025.

[53]World Health Organization. Ethics and Governance of Artificial Intelligence for Health: Guidance on Large Multimodal Models[R]. Geneva: World Health Organization, 2024.

[54]UNESCO. Guidance for Generative AI in Education and Research[Z]. 2023.

[55]伍雪莲,王丽桃,曾恩泉. 生成式人工智能在医疗决策中的伦理困境与治理路径探析[J]. 科学咨询,2026(10):17-21.

[56]弓孟春,马永慧,潘慧,等. 生成式医学人工智能临床应用的伦理治理专家共识(2025 年版)[J]. 中国医学科学院学报,2026,48(1):24-37.

[57]弓孟春,李雨杭,马永慧,等. 生成式医学人工智能的伦理治理:三维协同路径与中国实践[J]. 医学信息学杂志,2026,47(1):2-8,23.

[58]朱雯熙,闫紫柔. 德国医学领域生成式人工智能伦理治理路径解析及启示[J]. 医学与哲学,2025,46(19):27-31.

[59]邹文卿,沈昊创. ChatGPT 在医疗领域中的应用、伦理风险与治理研究[J]. 医学与哲学,2023,44(21):7-11.

A Review of Artificial Intelligence Literacy Education for Medical Students in the Context of Generative Artificial Intelligence from the Perspective of Teaching Evaluation Reconstruction

XU Ming-chao, XIE Li-dan, XIE Qi-long

(Traditional Chinese Medicine Hospital of Meishan, Meishan Sichuan 620000, China)

Abstract: Objective: To review research progress on artificial intelligence literacy education for medical students in the context of generative artificial intelligence and to analyze its application scenarios and educational implications in medical teaching evaluation. Methods: A scoping review approach was used to retrieve Chinese and English literature from CNKI, Wanfang, VIP, PubMed/Europe PMC, OpenAlex and Crossref, as well as relevant documents on artificial intelligence governance. Results: A total of 59 publications were included, including 25 studies on teaching practice or evaluation. Generative artificial intelligence has entered question generation, case analysis, inquiry training, reflective writing assessment, medical record evaluation, virtual standardized patients and course evaluation. Conclusion: Teaching evaluation scenarios can present medical students' knowledge, skills and attitudes in using generative artificial intelligence, and may serve as an entry point for embedding artificial intelligence literacy education into curriculum, teaching, assessment and governance.

Key words: generative artificial intelligence; medical students; artificial intelligence literacy; medical education; teaching evaluation