·临床教学·

人工智能在临床推理教学新模式中的研究及展望

徐赫　焦洋

中国医学科学院/北京协和医学院/北京协和医院全科医学科(普通内科),北京 100730

通信作者:焦洋,Email:peterpumch@163.com

【摘要】 人工智能(artificial intelligence,AI)技术的发展正在驱动医学教育模式转型,其中临床推理(clinical reasoning)教学领域的变革尤为显著。传统临床推理教学模式依赖资深医师通过真实案例进行经验传授,而新技术的涌现使AI逐渐成为教学过程中的关键要素。本研究综述并分析近5年相关研究发现,AI技术通过4个核心要素构建临床推理教学新路径:基于AI的智能案例系统(用于材料准备)、虚拟病人教学系统(推理能力训练)、智能评分系统(实现能力评估),以及全程嵌入的AI决策支持系统。循证研究表明,AI在临床案例整合和真实情景模拟方面展现出优势,但其在提供临床推理信息的可靠性、时效性以及对学生临床推理能力长期影响方面仍需进一步验证。现阶段实践建议采用传统教学与新路径相结合的双轨制教学模式,在关键决策点保留人工审核环节,以有效控制AI可能带来的错误风险;在未来,具备临床推理能力的AI模型有望为医学教育提供更为可靠的支持。

【关键词】 人工智能;　临床推理;　医学教育

【中图分类号】 R-4

基金项目:中央高水平医院临床科研专项(2022-PUMCH-A-017);北京协和医院教育创新与改革项目(CMB-JYCXGG04)

DOI: 10.3760/cma.j.cn116021-20250310-02066

Research progress and prospects of artificial intelligence in new models of clinical reasoning education

Xu He, Jiao Yang

Department of General Internal Medicine, Peking Union Medical College Hospital, Chinese Academy of Medical Science & Peking Union Medical College, Beijing 100730, China

Corresponding author: Jiao Yang, Email: peterpumch@163.com

【**Abstract**】 The advancement of artificial intelligence (AI) technology has initiated a paradigm shift in medical education, particularly transforming the field of clinical reasoning training. While traditional training models rely on experienced physicians imparting knowledge through real clinical cases, the new technological revolution empowers AI to play an increasingly significant role in educational processes. Through reviewing and analyzing research data from the past five years on AI-assisted clinical reasoning education, this paper reveals that AI technology establishes a novel clinical reasoning education pathway through four core modules: AI-based intelligent case system (material preparation), virtual patient educational system (reasoning skill development), intelligent scoring system (competency assessment), and AI decision support system integrated throughout the process. Evidence-based research demonstrates the advantages of AI in clinical case integration and teaching based on realistic scenario simulation. However, the reliability and timeliness of AI-provided clinical reasoning information and its long-term impacts on students' clinical reasoning capabilities require further verification. Currently, a dual-track teaching model that integrates

traditional and novel approaches is recommended, with human oversight maintained at critical decision points to effectively mitigate potential risks of decision errors posed by AI. In the future, AI models with clinical reasoning capabilities hold promise for providing more reliable support in medical education.

【Key words】 Artificial intelligence;　Clinical reasoning;　Medical education

Fund program: Central High Quality Hospital Clinical Research Program (2022-PUMCH-A-017); Education Innovation and Reform Program of Peking Union Medical College Hospital (CMB-JYCXGG04)

DOI: 10.3760/cma.j.cn116021-20250310-02066

人工智能(artificial intelligence,AI)的快速发展正深刻地重塑着临床医学的实践范式。随着诸如DeepSeek[1]、Gemini[2]等具备推理能力的生成式大模型问世,AI在应对复杂推理任务时的效能与稳定性已显著提升。目前,AI技术已被诸多医疗机构纳入日常临床工作之中,承担起临床工作质量监控[3]、辅助临床决策[4]乃至虚拟医生角色[5]等重要职责。这一技术变革的浪潮给临床医生带来了全新的挑战,并催生出医学教育领域的一个核心议题:如何系统地培养新一代医生与AI协同工作的能力?对这一议题的深入探索,不仅关乎医学教育的具体实施,更直接关系未来临床工作质量的提升与医疗安全体系的构建。

作为临床医生必备的核心技能,临床推理(clinical reasoning)在AI时代的临床教学中正凸显出前所未有的价值[6]。该能力涉及复杂的认知过程:首要步骤是搜集相关的症状、体征等临床信息,随后借助影像学和实验室检查等手段,并结合患者特有的危险因素,来构建疾病诊断的初步假设。在此基础上,进一步形成临床处理方案,并通过实践验证或否定这些假设[7-8]。随着AI技术融入医学实践与教学,基于AI的临床推理教学新模式正经历快速演化,这种新模式呈现出哪些具体形态?相较于传统模式有哪些优势与缺陷?当前尚未有研究对这一问题进行归纳与分析。因此,本文拟通过归纳并分析相关文献,从AI时代下医学临床核心能力培养的重点出发,系统解析AI辅助临床推理教学培养的新途径,为未来构建临床医生与AI协同工作提供理论支撑与实践指引。

1　人工智能浪潮下临床实践变革

AI技术在临床医学领域展现出显著应用潜力。其中,判别式模型可以通过电子病历数据实现疾病诊断预测,其分类诊断的准确性已与临床医生相当[9-10];生成式大语言模型则在生成临床推理思路、诊疗计划等方面展现较大潜力,研究显示其在标准化临床案例分析中的表现可超出初级住院医生的水平[11]。随着更强大的生成式模型的出现,AI可以辅助从图像识别到临床诊疗等一系列工作[12],这不仅为医疗实践带来新的变革,更对临床工作者的核心能力提出新的挑战。

尽管大语言模型在上述测试任务中展现出优势,但其临床实用化进程仍面临困境。现有证据表明,一方面,当前模型虽在分析现实病历时的诊断正确率、临床推理准确性与主治医师、住院医师相当,但是在回答临床专科领域问题方面的表现仍不如专科医生[10],且会生成更高频次的错误临床推理假设[13];更为关键的是,随机对照试验提示AI辅助临床推理存在效率悖论,即采用大语言模型辅助的医生在进行决策推理时并未能提高准确率和速度[14]。这些发现提示,在临床推理中使用AI不仅未必能提高决策准确性,还有可能对医生临床推理能力发展有潜在的负面影响。尽管相关研究尚处于初级阶段,但现有证据已揭示,在临床场景中整合AI仍存在医学科学、医学教育、医学伦理和政策管理等多个方面的困难,需要多方专业人士的共同关注和突破[15]。

值得注意的是,随着AI大语言模型以及相关系统的大量涌现,医学生以及低年资医生已经能够迅速获取海量临床推理所需的信息,但同时也需应对信息来源不明确、临床推断逻辑有误等挑战,AI编造事实的"幻觉现象"[16]仍频发,这要求医学生具备识别并自主推理判断的能力,从而避免AI误导临床决策的情况发生。

2　人工智能时代临床核心能力的培养目标

AI在医疗领域的核心功能应定位在辅助提升医疗服务质量,而非取代临床工作者的专业判断与

监督职能。那么，面向未来且拥抱AI的医学教学目标则为培养具备独立临床判断能力且能与AI协同工作的临床工作者。该教学目标要求临床工作者实现两方面能力提升：一方面需熟练操作AI系统以提升信息获取效率；另一方面具备对AI生成信息的批判性鉴别能力，并在此基础上发展独立临床推理能力。如何培养这种高效获取信息和独立推理判断的能力呢？一方面，这种能力来自对AI临床决策支持系统本身的深度认知。例如，来自巴黎大学医学院的团队设计了一项AI教学课程——人工智能和临床决策支持系统（Artificial Intelligence and Clinical Decision Support Systems，AI-CDSS）[17]，课程中医学生采用AI-CDSS的设计者视角展开学习，这种角色转换训练可以帮助医学生更好地适应智能化医疗系统，训练批判性使用数字医疗服务的能力。

另一方面，需更加注重医学生自主临床推理能力的系统性培养。临床推理能力在医学教学中面临挑战，主要源于其隐性知识属性：临床推理过程往往基于经验积累而难以外显化。传统医学教学主要采用学徒制模式，通常由临床经验丰富的教授通过真实案例展开教学；教授在呈现复杂的临床推理思维过程时，会有意识地将相对应的临床事实与思考过程进行结构化呈现。这种基于案例的学习模式（case-based learning，CBL）通过多个相似案例的对比分析[18]，有效强化了临床推理能力的训练效果[19]。

值得注意的是，AI为临床推理能力教学带来挑战的同时也带来了机遇。新冠疫情的全球蔓延加速了AI与医学教育的融合进程[20]：以纽约大学医学院为例，该校在2020年率先推出精准医学教育模式（precision medical education，PME）[21]，通过全程AI辅助教育，在个性化基础医学课程和临床推理训练方面均取得显著成效。在此背景下，2022年初生成式大语言模型ChatGPT公开发布，进而有研究证实其可以通过美国医师执照考试（United States Medical Licensing Examination，USMLE）[22]，该结果不仅促使医学教育领域反思AI时代的医学教育定位，更进一步推动了AI+医学临床推理能力教育新模式的产生[23-24]。

3　人工智能化医学临床推理能力教育的新模式

随着AI技术的突破性进展，一种新型临床推理教育模式正在形成（图1），该模式有望打破临床思维教学难以外显化的固有困境。与传统整合的教学路径不同，新型智能化教学路径基于AI整合的临床案例库，通过虚拟病人呈现仿真案例，最后采用智能评分系统评估学生学习成果；在此过程中，学生可借助AI决策支持系统获得临床推理所需信息。当前模式的技术实现路径与教学效能验证已成为医学教育研究的前沿领域[25]。本部分基于图1所示人工智能化教学框架，对AI在临床推理能力培养中的具体应用及实施效果展开探讨。

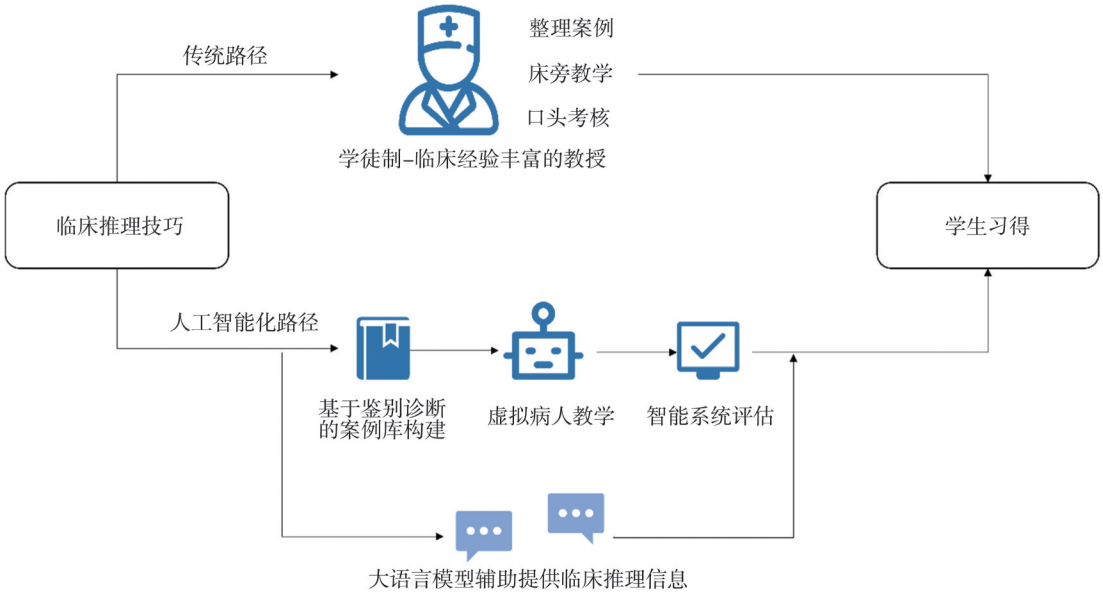


图1　技术变迁下临床推理能力教学的传统路径与人工智能化路径的对比

3.1　教学材料准备：基于AI的智能案例系统

传统临床教学因依赖高年资医师人工编制的教学案例，面临场景覆盖有限、病种单一等问题。在实际临床实践中，表现出相似症状群的患者可能对应着截然不同的诊疗方案，因此低年资临床医生需要大量临床推理过程类似、而推理结果不同的案例进行推理训练。AI技术可以将大量病历信息的整合并重组为上述案例材料，有效支撑临床思维训练。

典型范例包括Zack团队采用AI模型开发的临床推理过程编码病例库[26]。该研究团队将临床推理过程分为3个步骤：案例呈现、鉴别诊断和最终诊断。研究者邀请临床专家对新英格兰医学期刊（New England Journal of Medicine，NEJM）从1965年至2020年的2 525个临床案例进行分析，采用机器学习（machine learning）中的无监督学习方法（unsupervised learning），利用数据驱动的自然语言处理技术对这些病例的鉴别诊断过程进行系统编码，从而构建出临床推理病例库。除此以外，国内学者也开发出针对临床思维的AI训练平台，平台结合医疗系统数据，以手机软件的形式为住院医师提供涉及各种系统的65种疾病问诊、查体、诊断练习[27]。另有研究团队基于电子病历系统真实病例开发出虚拟案例学习系统，该系统通过数据驱动实现案例与临床思维训练阶段的动态适配，使用者可查看原始问诊与查体信息[28]。

相较于传统案例库，基于AI构建的系统可以通过鉴别诊断过程进行数据驱动分类，在分析的基础上使得案例服务于临床思维训练的各个阶段。需特别指出的是，基于学术期刊案例构建的系统存在疑难病例偏倚风险，而依托真实病历的系统则需要注重保护病人隐私，谨防数据泄露、数据被篡改的风险。另外，当前证据尚不能充分证明这些案例库的临床教学效果，有待未来进一步评估。

3.2　推理能力训练：基于AI的虚拟病人教学系统

3.2.1　AI虚拟病人基本概念

虚拟病人模拟器（virtual patient simulators，VPSs）[29]在AI的辅助下可以通过自然语言交互模拟真实诊疗场景，让学生无须进入病房即完成传统技能的训练[30]，这种虚拟病人模拟器相当于临床推理训练的全托管程序。

典型范例包括Furlan团队于2021年开发出的Hepius虚拟患者模拟器[31-32]，该系统能够模拟真实的患者对话和互动情境，构建高仿真的临床训练环境。Hepius模拟器同时依赖智能辅导系统和自然语言处理两大技术，其中，智能辅导系统可以根据学生的特定学习需求动态调整教学内容，而自然语言处理方法则可以实现学生与虚拟病人之间的互动，支持学生完成病史采集、体格检查、实验室检查申请、产生临床诊断假设等流程，其中病史采集部分需要学生输入病史采集问题，算法将输入问题与教师先前提供的参考问题列表进行模式匹配，选出模式类似的参考问题。

除此以外，Borg团队同样将大语言模型融入虚拟病人系统，使得虚拟病人可以模拟风湿性疾病，而且该系统可以根据学生的知识水平来调整难度[33]。而Brügge团队则仅采用大语言模型和一系列提示词设计出虚拟病人交互反馈系统，支持学生完成医患互动并为其提供结构化的反馈[34]。由此可见，在大语言模型的辅助下，教学得以采用交互模拟工具提供更为仿真、可以灵活调整的临床场景。

3.2.2　AI虚拟病人效果验证

虚拟病人是否可以稳定、有效地为学生提供临床场景呢？Hepius开发者进行了两项小样本的有效性研究，结果表明虚拟病人能有效提升医学生医学知识掌握水平（医学知识选择题的得分提高）[32]。当结构化反馈结合进入虚拟病人系统时，学生可以更快地习得并掌握临床推理能力[34]。另一方面，从学生的主观体验而言，学生们认为使用虚拟病人比线性计算机程序更具真实性，在训练临床推理、沟通和情绪回应技巧方面的主观感受更佳[33]。

从当前结果可知，虚拟病人提供了基本可靠且真实可感的教学环境，学生练习临床推理思维可以同时进行沟通技巧的训练。但是这种虚拟系统也有不足之处：为了确保其稳定性，整个程序是按照高年资医师设计的理想临床推理过程搭建的，这种开发的流程较为繁琐，限制了教学案例的数量与还原度。另外，固定的临床推理思路限制了学生开放性思考的空间，潜在地限制了学生批判性思维的发展，这一争议点仍需后续研究深入探讨。

3.3　能力评估阶段：基于AI的临床推理评分系统

传统临床推理能力评估往往依赖高年资医生的口头考核或者文本评分，该过程人力成本较高。在AI尚未大规模发展时期，已经有研究者采用人工

神经网络进行教学评分[35]，随着近期自然语言处理技术的发展，已有 AI 评分系统出现在临床教育实践中。

例如，多个团队开发出 AI 自动评分系统[36-38]。这种评分系统针对的文本为病历文本，病历文本可以较好展示学生的临床推理思维。与其他文本类型的差别在于，病历是长度较短、内容不完整的文本类型，自动评分系统需结合自然语言处理和深度学习模型完成逐个短语的分析，进而将每个病历与评分表进行逐项对照。这种系统不仅能够快速处理大量病历记录，还能通过提供逐个短语对应的反馈标记帮助学生提高病历书写技巧。

对上述评分系统的验证结果表明，模型评分与人为评分具有较高一致性，表现出较好的信效度[36-38]。类似的评分系统可以减少学生对 AI 工具的依赖性倾向，促进其自主开展灵活的临床推理训练。

3.4 智能信息辅助：AI 决策支持系统

在整个学习过程中，大语言模型可根据学生需求生成临床推理文本。然而，大语言模型是否可以有效促进学习？Xie 等[39]选取了 3 个主流 AI 大语言模型（ChatGPT-4、Bard 和 BingAI），对这些模型辅助低年资医生进行临床决策的潜力进行了比较：研究者以低年资医生视角向模型输入临床情景的复杂问题，并对大语言模型给出的答案进行可靠性、准确性、信息的充分性、可读性等维度的评估，结果表明 ChatGPT-4 的表现最佳。然而值得注意的是，类似的研究表明 ChatGPT 虽在流行率、诊断和治疗等基础问题中表现尚可，但其对于进一步深入问询以及在临床推理方面表现较差，且有时会遗漏关键信息[23,40]。

以上证据表明，尽管大语言模型可以稳定输出临床推理文本，但其仍然无法规避错误以及其相应的临床风险[41]。有研究对比了 ChatGPT、网络资源库和临床指南的辅助下医学生分析案例的准确度和完整性，结果表明，临床指南辅助组的准确度和完整性均为最佳，ChatGPT 组次之[42]。该研究采用较小的样本量反映出大语言模型支持医学生完成临床推理任务虽然具备优势，但并非最佳选择。其他研究同样表明大语言模型的支持并不能使临床工作者更好地完成临床推理任务[14]，这提示在 AI 时代医学教育需重点培养错误信息甄别能力，否则可能加重临床工作认知负担与医疗风险。

4 人工智能在临床推理教学中的挑战与展望

4.1 人工智能在临床推理教学中的挑战

AI 正逐步深入地融入医学实践与医学教育。根据以上几个部分的分析，可以得知 AI 正推动临床推理教学模式变革，尽管当前技术尚未完全成熟，但现有技术已可以基本实现 AI 与传统路径相结合的双轨制教学方式[43]；将 AI 与传统 CBL 相结合，可以在拓宽教学维度的同时规避现有模型出错的风险。随着 AI 技术持续发展，传统教学过程进一步模块化、自动化，医学生有望实现自主临床案例分析和推理训练。这种自主学习的模式不仅降低了教育成本，还使得学生能够更加熟悉在 AI 系统中进行互动、在更灵活的环境中掌握复杂的临床技能。此外，AI 评分系统的出现为学生提供了及时、客观的反馈，提升其病历书写和病例分析的能力，提高了教学效率。

然而，目前 AI 在临床推理教学中的应用仍面临诸多挑战。首先，当前大语言模型输出内容依然存在一定比例的事实错误，在医学教育情境中，若大模型偶然生成不完整、不准确或违背伦理的内容，不仅可能损害教育质量，最终还将对医疗服务质量产生负面影响。其次，AI 大模型无法实时补充最新的学术信息，而实时训练新的大模型成本过高。对于依赖高年资医师的知识和经验构建的案例库，其规模、多样性和时效性同样受到限制，无法涵盖知识快速更新且复杂多变的临床场景。再次，AI 教学效果验证不足：现有研究多基于小样本测试，而 AI 系统的长期教学效果及其在真实教学环境中的应用仍需进一步验证。未来的研究尤其需要关注医学生对 AI 大模型的依赖性，对模型的过度依赖反而可能限制临床推理思维的发展。

4.2 人工智能在临床推理教学中应用的展望

展望未来，为确保医学教育的完整性，既要审慎评估 AI 大模型生成准确、无偏内容的技术可靠性，也要系统分析其在实际应用中的潜在局限。针对现有 AI 辅助教学路径的局限性，本研究提出以下几点解决路径：首先，鉴于当下的 AI 模型仍可能输出事实错误信息，医学教育者可以将图 1 所示的传统路径与人工智能化路径相结合，形成 AI 工具赋能的、资深教授监督引领下的教学模式，例如采用大

模型生成临床场景[44]、采用AI辅助的影像学工具[45]进入课堂。这种方式既可以避免错误信息对学生产生误导，也可以激发学生兴趣、引导其自主学习的能力。其次，更积极地发挥AI大模型在个性化学习的潜在优势，开发以学生为中心的医学教育模式。其中的一种方式为：通过智能系统评估学生在临床推理方面的认知特点、能力差异和学习偏好，动态调整教学策略。这种自适应系统能够结合以虚拟病人为例的智能互动工具，从而帮助学生吸收知识、全面提升教学效果，最终形成灵活自适应的教育体系。再次，推动基于临床推理的AI模型发展。无论是判别式模型还是生成式模型，其黑箱特性导致临床推理过程不可追溯，因而可能对资历尚浅的医学生造成误导。未来应提倡将AI模型与临床知识库、临床指南以及临床推理逻辑相结合，从而为包括医学教育在内的临床医学界提供可靠的信息辅助。

总之，AI的持续发展有望为医学教育提供技术支持，通过更为灵活的智能化教学体系，培养临床推理能力更强的医疗人才。

利益冲突　所有作者声明不存在利益冲突

作者贡献声明　徐赫：文献检索、数据整理、文章撰写；焦洋：问题提出、研究指导、论文审阅、经费支持

参考文献

[1] DeepSeek| 深度求索[EB/OL]. [2025-03-07]. https://www.deepseek.com/.

[2] Gemini-轻松对话，获取创意灵感[EB/OL]. [2025-03-07]. https://gemini.google.com.

[3] Zhang L, Yao L, Lu Z, et al. Current status of quality control in screening esophagogastroduodenoscopy and the emerging role of artificial intelligence [J]. Dig Endosc, 2024, 36(1): 5-15. DOI: 10.1111/den.14649.

[4] Gombolay GY, Silva A, Schrum M, et al. Effects of explainable artificial intelligence in neurology decision support [J]. Ann Clin Transl Neurol, 2024, 11(5): 1224-1235. DOI: 10.1002/acn3.52036.

[5] “AI医生”来了，未来将应用于哪些场景？_中央网络安全和信息化委员会办公室[EB/OL]. [2025-03-07]. https://www.cac.gov.cn/2025-02/25/c_1742183898096636.htm.

[6] Richards JB, Hayes MM, Schwartzstein RM. Teaching clinical reasoning and critical thinking: from cognitive theory to practical application [J]. Chest, 2020, 158(4): 1617-1628. DOI: 10.1016/j.chest.2020.05.525.

[7] Connor DM, Durning SJ, Rencic JJ. Clinical reasoning as a core competency [J]. Acad Med, 2020, 95(8): 1166-1171. DOI: 10.1097/acm.0000000000003027.

[8] Young M, Thomas A, Gordon D, et al. The terminology of clinical reasoning in health professions education: Implications and considerations [J]. Med Teach, 2019, 41(11): 1277-1284. DOI: 10.1080/0142159X.2019.1635686.

[9] Liang H, Tsui BY, Ni H, et al. Evaluation and accurate diagnoses of pediatric diseases using artificial intelligence [J]. Nat Med, 2019, 25(3): 433-438. DOI: 10.1038/s41591-018-0335-9.

[10] Shen J, Zhang CJP, Jiang B, et al. Artificial intelligence versus clinicians in disease diagnosis: systematic review [J]. JMIR Med Inform, 2019, 7(3): e10010. DOI: 10.2196/10010.

[11] Kayacan Erdoğan E, Babaoğlu H. Clinical reasoning and knowledge assessment of rheumatology residents compared to AI models: a pilot study [J]. J Clin Med, 2024, 13(23): 7405. DOI: 10.3390/jcm13237405.

[12] Chang C, Shi W, Wang Y, et al. The path from task-specific to general purpose artificial intelligence for medical diagnostics: a bibliometric analysis [J]. Comput Biol Med, 2024, 172: 108258. DOI: 10.1016/j.compbiomed.2024.108258.

[13] Cabral S, Restrepo D, Kanjee Z, et al. Clinical reasoning of a generative artificial intelligence model compared with physicians [J]. JAMA Intern Med, 2024, 184(5): 581-583. DOI: 10.1001/jamainternmed.2024.0295.

[14] Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial [J]. JAMA Netw Open, 2024, 7(10): e2440969. DOI: 10.1001/jamanetworkopen.2024.40969.

[15] 江莉. AI医疗之“术”与“道”：基于道家思想的伦理反思[J]. 中国医学伦理学，2025，38(2)：201-206.

[16] Koenecke A, Choi ASG, Mei KX, et al. Careless whisper: speech-to-text hallucination harms[C/OL]//Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency. New York, NY, USA: Association for Computing Machinery, 2024: 1672-1681[2024-12-11]. https://dl.acm.org/doi/10.1145/3630106.3658996. DOI:10.1145/3630106.3658996.

[17] Tsopra R, Peiffer-Smadja N, Charlier C, et al. Putting undergraduate medical students in AI-CDSS designers' shoes: an innovative teaching method to develop digital health critical thinking [J]. Int J Med Inform, 2023, 171: 104980. DOI: 10.1016/j.ijmedinf.2022.104980.

[18] Thistlethwaite JE, Davies D, Ekeocha S, et al. The effectiveness of case-based learning in health professional education. A BEME systematic review: BEME Guide No. 23 [J]. Med Teach, 2012, 34(6): e421-e444. DOI: 10.3109/0142159X.2012.680939.

[19] Lessing JN, Pierce RG, Dhaliwal G. Teaching more about less: preparing clinicians for practice [J]. Am J Med, 2022, 135(6): 673-675. DOI: 10.1016/j.amjmed.2022.01.060.

[20] Lucey CR, Johnston SC. The transformational effects of COVID-19 on medical education [J]. JAMA, 2020, 324(11): 1033-1034. DOI: 10.1001/jama.2020.14136.

[21] Triola Marc M, Jesse B. Precision medical education [J]. Acad Med

J Assoc Am Med Coll, 2023, 98(7): 775-781. DOI: 10.1097/ACM.0000000000005227.

[22] Kung TH, Cheatham M, Medenilla A, et al. Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models [J]. PLoS Digit Health, 2023, 2(2): e0000198. DOI: 10.1371/journal.pdig.0000198.

[23] Sauder M, Tritsch T, Rajput V, et al. Exploring generative artificial intelligence-assisted medical education: assessing case-based learning for medical students [J]. Cureus, 2024, 16(1): e51961. DOI: 10.7759/cureus.51961.

[24] Mbakwe AB, Lourentzou I, Celi LA, et al. ChatGPT passing USMLE shines a spotlight on the flaws of medical education [J]. PLoS Digit Health, 2023, 2(2): e0000205. DOI: 10.1371/journal.pdig.0000205.

[25] Gordon M, Daniel M, Ajiboye A, et al. A scoping review of artificial intelligence in medical education: BEME Guide No. 84 [J]. Med Teach, 2024, 46(4): 446-470. DOI: 10.1080/0142159X.2024.2314198.

[26] Zack T, Dhaliwal G, Geha R, et al. A clinical reasoning-encoded case library developed through natural language processing [J]. J Gen Intern Med, 2023, 38(1): 5-11. DOI: 10.1007/s11606-022-07758-0.

[27] Wang H, Zheng D, Wang M, et al. Artificial intelligence-powered training database for clinical thinking: app development study [J]. JMIR Form Res, 2025, 9: e58426. DOI: 10.2196/58426.

[28] Wang M, Sun Z, Jia M, et al. Intelligent virtual case learning system based on real medical records and natural language processing [J]. BMC Med Inform Decis Mak, 2022, 22(1): 60. DOI: 10.1186/s12911-022-01797-7.

[29] Posel N, McGee JB, Fleiszer DM. Twelve tips to support the development of clinical reasoning skills using virtual patient cases [J]. Med Teach, 2015, 37(9): 813-818. DOI: 10.3109/0142159X.2014.993951.

[30] Berman NB, Durning SJ, Fischer MR, et al. The role for virtual patients in the future of medical education [J]. Acad Med, 2016, 91(9): 1217-1222. DOI: 10.1097/acm.0000000000001146.

[31] Furlan R, Gatti M, Mene R, et al. Learning analytics applied to clinical diagnostic reasoning using a natural language processing-based virtual patient simulator: case study [J]. JMIR Med Educ, 2022, 8(1): e24372. DOI: 10.2196/24372.

[32] Furlan R, Gatti M, Menè R, et al. A natural language processing-based virtual patient simulator and intelligent tutoring system for the clinical diagnostic process: simulator development and case study [J]. JMIR Med Inform, 2021, 9(4): e24073. DOI: 10.2196/24073.

[33] Borg A, Georg C, Jobs B, et al. Virtual patient simulations using social robotics combined with large language models for clinical reasoning training in medical education: mixed methods study [J]. J Med Internet Res, 2025, 27: e63312. DOI: 10.2196/63312.

[34] Brügge E, Ricchizzi S, Arenbeck M, et al. Large language models improve clinical decision making of medical students through patient simulation and structured feedback: a randomized controlled trial [J]. BMC Med Educ, 2024, 24(1): 1391. DOI: 10.1186/s12909-024-06399-7.

[35] Stevens RH, Najafi K. Can artificial neural networks provide an "expert's" view of medical students performances on computer based simulations [J]. Proc Annu Symp Comput Appl Med Care, 1992: 179-183.

[36] Bond WF, Zhou J, Bhat S, et al. Automated patient note grading: examining scoring reliability and feasibility [J]. Acad Med, 2023, 98(11S): S90-S97. DOI: 10.1097/ACM.0000000000005357.

[37] Cianciolo AT, LaVoie N, Parker J. Machine scoring of medical students' written clinical reasoning: initial validity evidence [J]. Acad Med, 2021, 96(7): 1026-1035. DOI: 10.1097/ACM.0000000000004010.

[38] Schaye V, Guzman B, Burk-Rafel J, et al. Development and validation of a machine learning model for automated assessment of resident clinical reasoning documentation [J]. J Gen Intern Med, 2022, 37(9): 2230-2238. DOI: 10.1007/s11606-022-07526-0.

[39] Xie Y, Seth I, Hunter-Smith DJ, et al. Investigating the impact of innovative AI chatbot on post-pandemic medical education and clinical assistance: a comprehensive analysis [J]. ANZ J Surg, 2024, 94(1/2): 68-77. DOI: 10.1111/ans.18666.

[40] Arbel Y, Gimmon Y, Shmueli L. Evaluating the potential of large language models for vestibular rehabilitation education: a comparison of ChatGPT, google gemini, and clinicians [J]. Phys Ther, 2025, 105(4): pzaf010. DOI: 10.1093/ptj/pzaf010.

[41] Maitland A, Fowkes R, Maitland S. Can ChatGPT pass the MRCP (UK) written examinations Analysis of performance and errors using a clinical decision-reasoning framework [J]. BMJ Open, 2024, 14(3): e080558. DOI: 10.1136/bmjopen-2023-080558.

[42] Montagna M, Chiabrando F, De Lorenzo R, et al. Impact of clinical decision support systems on medical students' case-solving performance: comparison study with a focus group [J]. JMIR Med Educ, 2025, 11: e55709. DOI: 10.2196/55709.

[43] Gim H, Cook B, Le J, et al. Large language model-supported interactive case-based learning: a pilot study [J]. Intern Med J, 2025. DOI: 10.1111/imj.70030.

[44] Zheng K, Shen Z, Chen Z, et al. Application of AI-empowered scenario-based simulation teaching mode in cardiovascular disease education [J]. BMC Med Educ, 2024, 24(1): 1003. DOI: 10.1186/s12909-024-05977-z.

[45] Zhang Y, Feng H, Zhao Y, et al. Exploring the application of the artificial-intelligence-integrated platform 3D slicer in medical imaging education [J]. Diagnostics (Basel), 2024, 14(2): 146. DOI: 10.3390/diagnostics14020146.

（收稿日期：2025-03-10）
（本文编辑：兰明娟）