

# 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

陆璐,丁峰,肖顺华  
(江苏卫生健康职业学院,江苏南京 211800)

**[摘要]**人工智能(Artificial Intelligence, AI)技术的快速发展,给教育带来了革命性影响,尤其是以 ChatGPT 为代表的生成式人工智能(Generative Artificial Intelligence, GenAI)更是掀起了教育变革的新浪潮。生成式人工智能技术嵌入医学教育教学评价,可以有效丰富评价内容、完善评价指标、优化评价方法、开展个性化教学,但也带来了考试诚信存在风险、信息准确性缺乏有效甄别、医学人文精神培养孱弱、数据安全和个人隐私存在泄露风险等众多挑战。对此,广大医学教育者应从丰富考试形式、加强医学生批评精神和医学人文精神的培养、强化数据安全和隐私保护等方面推动人工智能时代医学教育教学评价新形态的高质量发展。

**[关键词]**生成式人工智能;医学教育;教学评价

**[中图分类号]** G642.47 **[文献标识码]** A  
**doi:**10.3969/j.issn.2096-711X.2025.10.053

**[文章编号]**2096-711X(2025)10-0158-03

**[本刊网址]** <http://www.hbxb.net>

2022年底,美国人工智能研究室 OpenAI 发布 ChatGPT,因其出色的对话和理解能力,迅速引发了全球界的关注。紧接着,2023年3月,OpenAI 更是在全面升级 GPT-3 的基础上发布了多模态预训练大模型 GPT-4,后者更具有创造性和协作性,更是引起了现象级的火爆。随后,百度的文心一言、讯飞的星火认知、阿里云的通义千问、腾讯的混元大模型等相继面世,生成式人工智能这股热潮更是达到了前所未有的巅峰,对教育领域也带来了革命性的影响。

医学教育是培养医疗专业人才的重要方式,而教学评价则是检验、提升教学质量的关键手段。传统的医学教育教学评价主要以笔试、实践操作、撰写报告等形式开展,存在诸多局限性。随着生成式人工智能的发展,AI 赋能教学改革成为广大教育者研究的热点。但是,当我们聚焦于新兴技术带来的发展机遇的同时,也需要关注其内隐性的风险,尤其是在医学教育教学评价中,生成式人工智能带来的到底是机遇还是挑战?这成为广大医学教育者关注的重要议题。

## 一、生成式人工智能理论概述

### (一)生成式人工智能的内涵

生成式人工智能是人工智能的一个分支,是基于深度学习模型的算法,可通过学习数据分布的规律,自主生成与原始数据类似的内容。初期,生成式人工智能只能生成丰富的文字信息,现已逐步发展为能够生成更多的形式,如图片、声音、视频、代码等,实现输出风格多样化。

### (二)生成式人工智能的特征

作为一种新兴的科技生产力,生成式人工智能是在集合计算机科学、逻辑学、哲学、心理学等众多学科的基础上发展出来的新技术,具有以下四个典型特征:①交互性:通过人机互动,用户只需要在对话框内输入指令或问题,即可从聊天机器人那里得到高质量的结果,包括文档编辑、文案设计、视频生成等。这种交互性符合人类的交流习惯,易于操作,便于上手。②“类人”性:与传统的搜索引擎不同,生成式人工智能输出的结果并不是从网络中直接搜索而得,而是通过自我学习后对知识进行重组,进而根据用户指定输出逻辑和语

法都合理的创新性的内容。也就是说,生成式人工智能能够模拟人类思考并进行独立创造,这正是其最大的优点,也是其产生巨大影响的关键所在。③个性化:即根据用户需求生成作品,只要用户输入的指令不同,生成式人工智能输出的结果就不相同。④潜在风险性:生成式人工智能生成的作品缺乏监管,可能存在错误或诱导,甚至有时会“一本正经地胡说八道”,需要用户进行信息筛选和价值判断。

## 二、生成式人工智能应用于医学教育教学评价的优势分析

### (一)有利于丰富考试形式

随着“新医科”的发展,知识的简单传授和学生的浅表学习已难以适应医疗健康的需求,以患者为中心、以系统为基础、基于胜任力与结果导向,强调学生转化式学习的第三代医学教育改革随之开启。生成式人工智能引入教学评价,可以促进传统的固定式的笔试形式拓展为开放式的、以解决实际临床问题为中心的考核形式。比如,可将传统的由选择、名词解释、填空等简单题型组成的试卷转变为由临床实际问题所构成的综合型任务;将学生独立闭卷答题转变为基于“人机+人际”互动协作的团队合作;将标准化答案转变为非标准答案或实际临床解决方案,从而更好地考查学生的临床综合能力和职业素养。

### (二)有利于拓展评价指标,提高评价结果的准确性

人工智能技术基于大数据和算法模型,能实时跟踪和记录学生分析问题、解决问题的过程。因此,教师可以利用过程性数据,构建包含在线学习行为、思考过程、问题解决能力、团队协作等方面的综合性评价指标,改变原先的“唯分数论”或“一考定结果”的考试现状,对学生的学习过程和学习成果进行多维度、客观的评价。教师还可以引入增值评价理念,借助大数据智能分析系统开展学生在一个学习阶段内知识、能力、素质上的增值表现,以综合性、长周期的评价服务学生的学习发展。

### (三)有利于开展虚拟病例和情景模拟考试

在医学专业课程考核过程中,病例讨论、案例分析是常见的考核方法。随着计算机科学的发展,先进模拟技术被逐

收稿日期:2024-10-14

基金项目:本文系江苏省教育科学“十四五”规划课题青年项目“数字赋能视域下医学职业教育增值评价研究与实践”阶段性成果(项目编号:C/2023/02/26);江苏卫生健康职业学院院级在线开放课程重点项目研究成果(项目编号:JSJKK201706)。

作者简介:陆璐(1984—),女,江苏南通人,江苏卫生健康职业学院副教授,主要从事高职医学职业教育、教学评价研究。

渐引入模拟医学教育。一方面,生成式人工智能可以根据需求进行情景模拟并生成各种虚拟病历,如急诊、术前术后管理、随访等不同情境下的医疗案例,学生可以在模拟情境中与虚拟对象产生实时互动,模拟病例分析、病理诊断、用药指导、治疗方案决策等活动,以帮助学生早期接触临床、熟悉诊疗流程。另一方面,人工智能还可以通过模拟临床专家的回答,引导医学生进行临床思考,并在有效规避实际风险的情况下指引学生大胆练习决策方案,从而提高学生临床思维和实践能力。

此外,对于强调医学生临床实践动手能力的医学教育而言,生成式人工智能工具还可以为医学生生成三维立体的人体模型和虚拟手术模拟器,帮助学生更好地认识和理解人体结构,并在规避临床风险的情况开展虚拟手术操作,锻炼学生的动手能力。

#### (四)有利于即时反馈评价结果,开展个性化教学

生成式人工智能的应用为实时监测和反馈提供了有力的支持。通过智能评价平台可以收集学生的学习个性化智能肖像,包括作业成绩、考试分数、课堂表现、在线学习行为及轨迹,实时生成个人学习效果评价和分析,帮助教师及时了解学生的学习状态、学习规律及存在的问题,从而根据学生的基础水平、兴趣爱好、学习能力,提供定制化的学习计划和辅导,推送个性化学习资源,真正实现因材施教。比如科大讯飞研发的“讯飞智医助手”机器人,不仅涵盖国家医师资格考试模拟真题,还能在数据分析的基础上根据学生的学习进度和掌握情况,进行有针对性的指导和示范,同时提供实时反馈,帮助学生及时发现并纠正操作中的错误。

总之,生成式人工智能技术的飞速发展,已然成为推动医学教育教学评价深入改革的重要驱动力。然而,人工智能是把双刃剑,这一技术革新的背后,医学教学评价也正面临着前所未有的挑战。

### 三、生成式人工智能时代,医学教育教学评价面临的挑战

#### (一)考试诚信面临风险

生成式人工智能的强大生成能力为医学生提供了便捷的信息获取途径,但同时也为考试诚信问题埋下了隐患。在过程性考核过程中,当遇到撰写学术论文、制作项目汇报 PPT 等考核内容时,学生可以利用人工智能工具快速进行信息查询、文献检索、内容归纳甚至论文初稿整理等,大大提高学习效率。但在便捷的同时,也有学生将人工智能生成内容直接或稍加改动后作为自己的成果提交,以此逃避学习和思考的过程。这种行为不仅破坏了考试的公平性和有效性,降低了教育评价本身所具有的强化和检验学习效果的作用,更严重削弱了学生学习的主动性,长此以往导致学生对智能技术产生依赖、失去探索精神,与医学教育的创新性和可持续性的培养要求背道而驰。

同时,人工智能成的学术内容未经充分审查和验证,直接用于作业、论文或研究报告,将严重损害学术研究的权威性并引发深刻的知识产权、科研诚信和学术伦理问题。

#### (二)信息准确性缺乏有效甄别

生成式人工智能本质上就是汲取海量零碎数据,通过深度学习算法形成与用户提问最相关的答案,其训练数据是基于互联网上的文本和信息,有些数据可能是错误的、过时的或不准确的,导致最终输出的内容存在一定的误导性。尤其是医学知识具有较强的专业性,在面对复杂病例时,人工智能仅能给出一些“泛泛而谈”的建议,无法结合患者的实际情况制定详细诊疗方案,甚至有时会“凭空捏造”,给出错误建议。而医学生因为专业知识储备不足,无法对人工智能生成内容进行有效甄别,常常被人工智能给出的诊疗建议迷惑和误导,从而阻碍其临床思维和实践能力的发展。

#### (三)医学人文精神培养薄弱

在临床医学教学评价中,常常会使用标准化病人来开展客观结构化临床考试,其中的标准化病人经过专业培训,能够真实、准确地模拟病人的症状、体征和病史等。当人工智能嵌入客观结构化临床考试后,冷冰冰的、虚拟化的人工智能机器人可以模拟真人进行语言沟通,但却无法模拟病人的体征和症状,医学生就无法进行很好的“医患共情”。而共情能力是医学生最重要的人文精神之一,缺乏“爱”的医疗,过度“程序化”“流水线式”的诊疗过程,必将严重影响医疗事业的健康发展。

此外,当医学生在学习过程中过度使用人工智能技术,极易减少师生互动、生生互动,久而久之,学生社会交往能力下降、群体意识淡化,对他人情感、意志表达感受模糊或迟钝,这都为今后的医患沟通埋下隐患。

#### (四)数据安全和个人隐私存在泄露风险

当前,生成式人工智能数据库缺乏监管机制,数据源的广泛使用可能涉及大量个人敏感数据,加剧个人信息隐私泄露的风险。同时,医学生在与人工智能机器人讨论病例或医学情景时,会有意或无意地泄露患者信息,导致某些敏感信息被系统识别并储存。此外,在使用智能系统进行学生能力评估、学习行为分析时,医学生的学习数据和记录也会被智能系统存储,比如医学生的学习过程、个性化学习路径甚至是个人的科研创新点。这些数据如果被一些不法分子窃取或用于不正当目的,就极易导致知识成果遭到窃取。因此,如何确保数据在收集、存储、分析及应用过程中的安全,是人工智能时代医学教育教学评价工作面临的一大挑战。

### 四、应对与建议

(一)加强评价体系顶层设计,丰富考试形式,强化考试诚信监管机制

作为医学教育领域最具代表性的评估理论之一,米勒金字塔学习评估理论将医学生由学习而获得的专业力能分为四个层级,即知道、理解、展示和实践。当前,尽管医学专业课程评价的形式日趋多元化,除了常规笔试之外增加了一些“现状调查”“综述书写”等考试形式,但这些考核层次仍然主要集中在“知道”和“理解”这两个层面,学生很容易简单粗暴地利用人工智能生成内容应付考核任务,久而久之对人工智能产生依赖。因此,广大教师必须对课程考核进行深化改革,在命题方面进行创新,提高“展示”和“实践”层面的考核占比。在考核内容上,多引导学生探究“是什么”“为什么”,改变当前“重知识”而“轻能力”的考核现状;在考核方法上,可以利用人工智能技术进行创新,设计项目研究、技能操作比赛、跨学科“医学+”作品设计等考核内容,多关注学生在一段时间内的学习增值表现;在考核主体上,可以通过学生自评、学生互评开展学生之间的相互监督,并借助智能终端设备、人工智能技术自动化反馈等教学辅助工具,进行作业的查重,加强考试诚信监管。

#### (二)加强批判精神的培养,预防人工智能依赖

要想用好人造智能这把“双刃剑”,使其更好的服务医学生的学。第一,从价值理念层面,教师要引导学生规范、正确使用生成式人工智能以服务自身学习。例如,通过举办专门的人工智能应用讲座或经验交流会,介绍如何使用人工智能相关工具、如何输入高质量的用户指定、使用人工智能辅助学习时哪些行为是被鼓励或禁止的等。第二,从技术层面,教师还可以给学生讲解生成式人工智能的训练和内容输出的运行方式,增强对其技术壁垒的理解,告诉学生因为人工智能目前在技术层面还缺少监督机制,导致其非常容易产生“AI幻觉”,这样学生就更能以批判性思维理性看待人工智能,也知道了人工智能只是辅助工具而非替代品。最后,从

操作层面,教师可以要求学生人工智能生成内容进行标记(如同标记参考文献一样),这样可以提高信息来源的透明度。由于有了清晰的来源标记,医学生就可以更加有针对性查阅相关资料去进行核实,以减少对人工智能的盲目依赖。

### (三)坚持“以人为本”,强化医学人文精神的培养

医学教育存在特殊性,技术手段无法代替临床实践,尤其是无法取代教师的言传身教以及在课堂、临床见习、实习教学过程中对医学生进行的人文素养教育。“生物—心理—社会”医学模式的提出,“以患者为中心”的医学人文关怀也已成为医生必备的职业素养。因此,医学教育者有义务和责任在人工智能赋能教学改革的过程中,进一步加强医学生的人文素质教育。比如,可以利用人工智能工具,通过互动式的教学方式,设立线上论坛或虚拟临床岗位体验,利用人工智能机器人作为24小时“助教”,供医学生分享想法、交流知识,让医学生身临其境地感受临床情境中的复杂性,引导学生建立对生命的敬畏之心。需要注意的是,在整个设计环节中,人工智能可以发挥“授业”“解惑”的作用,但“传道”的主体仍然是广大一线教师,教师要肩负起主要职责,通过教师的“润物细无声”,将医学生的信仰和信念根植于心。

### (四)强化数据安全与隐私保护

在利用人工智能技术进行教学评价的过程中,必须高度重视数据的安全性和隐私性保护。在技术层面,学校信息中心可以采用先进的数据加密技术,通过信息匿名化处理、数据加密、访问权限设置、定期审查等方式维护个人隐私信息安全,还需要建立完善的监控机制和日志记录系统,监控系统访问情况,及时发现异常操作。从理念层面,教师要对学生进行隐私保护等相关法律法规和伦理规范的普及,组织学生学《网络安全法》《数据安全法》和《个人信息保护法》,让学生认识到人工智能时代保护个人隐私的重要性,严格遵守医学教育领域的隐私保护法规,从源头防止数据被非法获取或用于不正当目的,保障师生及患者的隐私权益。

### 五、结语

在生成式人工智能技术快速发展的时代浪潮下,这股新质生产力正深刻影响着医学教育的方方面面,也促使着医学教育评价走向“数字化”“智能化”和“个性化”。然而,在这“相遇”又“相撞”的过程中,机遇总是与挑战并存,生成式人工智能技术在带来便利的同时,其特殊的“类人”特性也为医

学评价带来了不可忽视的风险。未来,在走好“人工智能+医学教育”发展的道路上,广大医学教育者应当以开放的态度拥抱新技术,直面挑战、发挥优势,积极推进融合人工智能技术的多元化评价体系改革,加强医学生批判精神和人文精神的培养,为健康中国背景下“新医科”的发展输送高质量的医学人才。

### 参考文献:

- [1]尚智丛,闫禹宏. ChatGPT 教育应用及其带来的变革与伦理挑战[J]. 东北师大学报(哲学社会科学版),2023(5):44-54.
- [2]王运武,周甜,邬亚楠. 人工智能创作物何以变革教育——基于 ChatGPT 类人工智能的理性思考[J]. 中国医学教育技术,2024,38(5):543-549.
- [3]詹希旎,李白杨,孙建军. 数智融合环境下 AIGC 的场域化应用与发展机遇[J]. 图书情报知识,2023,40(1):75-85.
- [4]万学红. 全球医学卫生教育专家委员会 21 世纪医学教育展望报告的启示[J]. 中国循证医学杂志,2011,11(5):477-478.
- [5]袁欢欢,全祉悦,林伟. 当医学课程考核遇上生成式人工智能:价值、挑战与未来展望[J]. 中国医学教育技术,2023,37(3):254-259.
- [6]刘晖,何欣悦,寇丽圆,任曼. 基于 CIPP 理论的 AI 赋能医学教学评价研究[J]. 中国医学教育技术,2024,38(5):600-605.
- [7]丛立先,李咏霖. 聊天机器人生成内容的版权风险及其治理[J]. 中国出版,2023(6):16-21.
- [8]冯苛,易超,丁玉. 客观结构化临床考试在临床医学专业医学生考核中的探索与实践[J]. 黄河科技学院学报,2024,26(8):97-100.
- [9]王启帆,于钦明. 人工智能嵌入医学教育的伦理风险及其应对[J]. 卫生软科学,2024,38(7):71-74.
- [10]佑镁,王旦,梁炜怡,等. ChatGPT 教育应用的伦理风险与规避进路[J]. 开放教育研究,2023(2):26-35.
- [11]范益,孙秀兰. 虚拟现实的时代:模拟医学教育的机遇与挑战[J]. 南京医科大学学报(社会科学版),2018(2):156-158.

## The Influence of Generative Artificial Intelligence Technology on the Evaluation of Medical Education and Suggestions for Countermeasures

LU Lu, DING Feng, XIAO Shun-hua

(Jiangsu Health Vocational College, Nanjing Jiangsu 211800, China)

**Abstract:** With the rapid development of artificial intelligence (AI) technology, it has brought revolutionary impact on education. In particular, the generative artificial intelligence represented by ChatGPT has set off a new wave of educational reform. Generative artificial intelligence (GenAI) technology can promote the reform of medical education teaching evaluation, such as enriching evaluation content, improving evaluation indicators, optimizing evaluation methods, and carrying out personalized teaching. However, behind this technological innovation, the medical education evaluation system is also facing many challenges, such as the risk of examination integrity, the lack of effective screening of information accuracy, the weak cultivation of medical humanistic spirit, and the risk of data security and personal privacy disclosure. Therefore, the majority of medical educators should promote the high-quality development of the new form of medical education teaching evaluation integrating artificial intelligence technology from the aspects of enriching examination forms, strengthening the cultivation of medical students' critical spirit and medical humanistic spirit, and strengthening data security and privacy protection.

**Key words:** generative artificial intelligence; medical education; teaching evaluation

(责任编辑:桂杉杉)