

医学研究生应用生成式 AI 的学术诚信风险识别与应对机制

张再球[#] 李思杰[#] 刘韬华 赵明一 文 敏^{*} 何庆南^{*}

中南大学湘雅三医院儿科 湖南 长沙 410000

摘 要: 生成式人工智能 (AI) 正在改变医学研究生查阅文献、处理数据和撰写论文的方式, 也使学术责任的界定更为复杂。现有讨论多强调工具带来的效率提升, 对医学科研各环节中风险如何产生、责任如何落实, 仍缺少连贯梳理。本文围绕文献检索与使用、数据分析和论文撰写三个场景, 分析虚假引文、分析过程不透明、偏见放大、替代性写作及科研能力弱化等问题, 并根据风险出现的具体环节提出核验、披露、数据保护和能力训练等要求。在此基础上, 文章将生成式 AI 的使用边界、披露义务、核查责任、数据与隐私保护及能力培养归纳为行为准则, 旨在为高校制定医学研究生 AI 使用规则、开展科研诚信教育提供可执行的参考。

关键词: 生成式人工智能; 医学研究生; 学术诚信; 行为规范

引言

教育部于 2018 年印发《高等学校人工智能创新行动计划》, 将科技创新、人才培养和国际合作列为高校发展人工智能的重要任务, 为新一代人工智能发展提供政策支撑^[1]。在医学研究生培养中, 这类技术的价值并不只体现在信息处理速度上。借助临床情境模拟和罕见病例训练, 学生能够在较低风险的环境中反复练习, 并根据不同学习需求获得更有针对性的内容, 由此提高学习和实践效率^[2]。

工具使用范围不断扩大, 也使科研责任的界限变得更难把握。医学研究生可能在检索、分析或写作环节借助 AI, 但结果的真实性、方法的适切性以及文字所承载的学术贡献, 仍须由研究者本人负责。由此产生的问题不只是如何判断某段文字是否由 AI 生成, 更在于哪些任务可以交由工具辅助、哪些决定必须由研究者独立完成。本文据此考察医学研究生使用生成式 AI 时可能出现的学术诚信风险, 并尝试从具体科研环节提出相应的约束与支持措施。

1 生成式人工智能在医学研究生教育中的发展现状

医学信息的数量和复杂程度持续上升, 传统教学已难以在有限时间内覆盖全部临床知识与训练情境。AI 因而逐步进入医学教育和临床教学, 并被用于信息整合、情境模拟及学习支持^[3]。相关应用目前仍处于探索阶段, 但已有研究显示, AI 驱动的虚拟患者或模拟平台可让学生练习诊断与治疗决策, 并在进入真实临床环境前获得反馈^[4]。在非面对面教学中, 生成式 AI 还可根据学习者

的提问提供差异化解释, 这对维持学习动机和提升自主学习能力具有一定帮助^[5]。自然语言处理工具则能协助梳理医学文献、定位研究进展和临床指南^[6]。这些用途提示, AI 更适合作为补充性的学习支持, 而不能取代临床经验、教师判断和系统的专业训练。

技术效用并不能消解伦理责任。一项刊于《柳叶刀》的研究审计了近 250 万篇生物医学论文, 发现伪造引文在两年内增加约 12 倍, 说明引文真实性已成为需要持续关注的问题^[7]。对处于科研训练阶段的医学研究生而言, AI 使用至少涉及三个层次的要求, 即守住求实诚信的底线、服务个人学术成长, 并最终指向负责任的知识创新^[8]。学校和医院可通过伦理课程、审查与问责机制以及必要的检测手段降低风险。我国已先后发布《生成式人工智能服务管理暂行办法》和《关于在科研活动中规范使用人工智能技术的诚信提醒》^[9,10], 为规范使用提供了基本依据。不过, 这些文件以原则性要求为主, 对医学研究中的患者数据、统计分析责任和论文写作边界尚缺少细化规则, 具体执行与监督机制仍有完善空间。

2 医学研究生应用生成式 AI 的场景分析与风险识别

要提出可落实的规则, 需先回到医学研究生实际使用 AI 的科研环节。下文分别考察文献检索与使用、数据分析和论文撰写, 并据此识别各环节中较易被忽略的诚信风险。

2.1 文献检索与使用

文献检索是医学研究生开展科研工作的基础, 亦是耗费大量时间与精力的关键环节。传统检索模式依赖研

研究者人工拟定关键词、逐库检索并逐篇筛选,不仅流程冗长,且对检索者的知识结构与检索策略要求较高。生成式 AI 技术的快速发展正在深刻改变这一局面:基于大语言模型的文献检索工具已能深度覆盖循证医学检索、语义搜索和文献推荐等多种功能,使研究者得以通过对话式交互描述研究需求,由 AI 自动解析问题、生成优化关键词并执行跨数据库检索,在大幅压缩时间成本的同时,有效突破了传统检索中查全率与查准率难以兼顾的瓶颈^[11]。

但同时与之相伴的是一类系统性风险——AI 的“幻觉”(hallucination)问题。研究表明,AI 生成的文本具有表面合理性与高度欺骗性,在文献指引上极易凭空虚构:GPT-3.5 生成的参考文献中高达 55% 为伪造条目,GPT-4 虽有改善,伪造率仍达 18%;即使在真实存在的引用中,GPT-3.5 和 GPT-4 分别有 43% 和 24% 存在作者、期刊、卷期页码等实质性错误^[12,13]。更为严峻的是,这一问题已从个案演变为系统性现象:一项针对 250 万篇生物医学论文的审计研究发现,AI 生成虚假引用的行为正以可观速度扩散^[14]。究其根源,通用大模型因缺乏与 PubMed、Web of Science 等权威学术数据库的实时强绑定与底层校验,其训练语料中充斥着海量非学术性网络文本,导致输出结果存在严重的“信息异质化”与低质干扰^[15]。医学研究生处于学术训练的关键阶段,若对 AI 检索结果缺乏批判性核查能力,虚假文献引用与低质信息干扰将同时侵蚀其学术诚信意识与科研产出的可靠性。

2.2 数据分析

数据分析是医学研究生科研工作的核心环节,也是生成式 AI 介入最深、争议最集中的领域之一。医学研究的数据分析涵盖从基础统计描述、假设检验到机器学习建模、生物信息学分析等多个层次,对研究者的统计素养与方法学训练提出了较高要求。生成式 AI 的介入,使得这一环节的“效率提升”与“风险积累”之间的张力尤为突出。

在应用层面,医学研究生已借助 ChatGPT、DeepSeek 等大语言模型完成数据清洗、统计方法选择、代码生成与结果解读等多项任务。研究者有效跨越了 R 或 Python 等传统编程语言的技术壁垒,通过对话即可输出可视化图表,并高效完成 t 检验、卡方检验及 Logistic 回归等推断性统计。已有系统性综述指出,AI 对医疗教育和研究的最大吸引力,正在于其能够降低数据分析的技术门槛,使不具备编程背景的临床研究者也能快速完成数据处理^[16]。在进阶分析层面,生成式 AI 已被用于医学影像特征提取、基因表达差异分析和临床预测模型构建等更复杂的任务;更有研究者已将 AI 工具整合到 Meta 分

析、系统综述的数据提取与效应量计算流程中,显著缩短了研究周期。

然而,生成式 AI 在数据分析中的应用也催生了三类相互交织的学术诚信风险。最为突出的是黑箱分析导致的结论失真与代码隐性缺陷。AI 执行统计分析时,其内部推理过程对用户是不透明的:模型可能选择不当的统计方法、忽略数据分布假设,或对缺失数据执行错误的填补处理,而统计素养不足的研究生难以识别这些隐蔽错误。尤其在代码生成场景中,这一风险被成倍放大:当不具备编程基础的医学研究生直接运行 AI 生成的代码时,极易陷入难以察觉代码中潜藏的逻辑漏洞(如变量错配、循环条件错误或算法包调用不当)。这些语法正确但逻辑谬误的代码会导致分析结果南辕北辙。已有研究明确警示,用户必须对 AI 在量化分析场景中的输出准确性保持高度警惕^[17]。更严重的是,AI 在生成统计结果时同样存在幻觉现象——输出看似合理但实际与数据无关的统计量或 P 值,此类错误若未被发现,将直接导致不可重复的研究发现,侵蚀生物医学研究的科学基础。

与之并存的还有数据解读中的偏见放大问题。AI 的训练数据本身可能包含系统性偏见,当研究生将 AI 用于结果解读和讨论撰写时,这种偏见可能被放大并嵌入最终研究报告。AI 倾向于生成符合用户预期的“阳性”结果解读,可能加剧学术界已存在的发表偏倚问题^[18]。从治理视角看,AI 对学术不端概念的冲击体现在原创性、错误性与公平性等多个层面,数据分析场景中的 AI 误用正在重塑学术不端的边界^[19]。

更为深层的风险在于,对 AI 分析能力的过度依赖正在侵蚀研究生的核心科研素养。当研究生习惯于将数据“丢给 AI”获取分析结果时,其对统计学原理、研究设计逻辑和批判性思维的训练将严重不足。使用者的学术诚信意识淡薄与追责制度宽松,是生成式 AI 引发学术诚信危机的深层原因^[20]。这种“能力空心化”风险在医学研究生教育中具有特殊严峻性:未来这些研究生将成为独立的研究者或临床决策者,若其数据分析能力完全建立在 AI 依赖之上,一旦脱离 AI 工具或遭遇 AI 出错的情形,将无法独立做出正确的科学判断。

2.3 论文撰写

论文撰写是医学研究生科研产出的最终环节,也是生成式 AI 应用最为广泛、争议最为集中的领域。从文字润色、摘要撰写到全文生成,AI 工具已深度介入学术写作的各个阶段。如何在提升写作效率与维护学术诚信之间划清边界,已成为医学研究生教育亟待解决的难题。

在辅助写作层面,生成式 AI 可以协助完成语法纠

错、句式优化、段落重组等基础文字处理任务,帮助非英语母语的研究生提升英文学术写作质量。调查显示,AI对于刚踏入学术领域的研究生具有正向溢出效应,能够帮助学生更好地理解 and 遵循学术规范,但知识产权归属与学术诚信的边界问题同样不容忽视^[21]。在中文学术写作场景中,AI还被广泛用于学位论文的文献综述撰写、研究方法描述与讨论部分扩写,成为研究生应对紧迫截稿期限的重要工具。

在文本生成层面,生成式AI的能力已从辅助润色扩展到可独立产出结构完整、逻辑连贯的学术段落甚至全文。多学科专家评估认为,AI在内容生成方面具有强大能力,但过度依赖AI生成文本将导致论文思想深度和独创性的结构性下降^[22]。已有研究将AI在学术写作中的使用划分为三个伦理层级:语法检查与翻译属于伦理上最可接受的方式;大纲生成和内容润色属于有条件可接受;而从零生成全文、数据解读和文献综述则属于伦理上可疑的范畴^[23]。

论文撰写环节中AI应用引发的学术诚信风险同样体现在三个维度。其中最具争议性的是AI代写与署名权问题。当研究生使用AI生成论文核心内容而未作实质性修改时,实质上构成了“AI代写”行为。AI生成文本给传统抄袭检测带来了根本性挑战:AI生成的文字在字面上是“原创”的,不构成传统意义上的逐字抄袭,但其思想内容和知识贡献并非来自署名作者^[24]。我国教育部已联合多所高校发布相关指导意见,明确将“完全或主要使用AI生成论文内容并作为个人成果提交”列为学术不端行为。传统以文字重合率为核心的检测标准,已不足以应对这类新型学术不端行为^[19]。

AI幻觉导致的虚假内容嵌入构成了另一重严峻风险。AI在科学写作中可能生成包含虚构数据、错误引用或逻辑谬误的文本,且这些内容在表面上与人类书写无异。最早系统描述这一现象的研究指出,ChatGPT可能编造看似合理但完全虚构的研究发现^[13]。后续综述进一步揭示:ChatGPT生成的参考文献中高达38%存在DOI错误或完全伪造,16%的文章完全虚构,近50%的参考文献存在不同程度的错误,仅7%完全真实且准确^[23]。同时有研究发现,研究人员可利用GPT-4在极短时间内生成包含数百名虚拟患者的伪造临床数据集,该数据集不仅各项特征极具迷惑性,甚至能通过常规统计学检验得出支持特定结论的“阳性结果”,传统同行评审难以察觉其真伪。^[25]对于医学研究生而言,若未对AI生成内容逐一核实,虚假信息可能被直接纳入学位论文或期刊投稿,其后果不仅是个人学术声誉的损害,更可能污染后续研究者的引用链条,造成学术生态的持续性损害。

此外,写作能力的结构性退化是一个更长期且不易察觉的风险。长期依赖AI进行论文撰写,可能导致研究生的学术写作能力和独立思考能力发生不可逆的结构性退化。调查显示,学生虽然普遍认可AI在写作辅助中的价值,但也对准确性、隐私及对个人发展的影响表示担忧^[15]。这种担忧具有现实依据:当研究生将AI视为“写作拐杖”,其构建学术论证、组织逻辑链条和凝练学术观点的能力将因缺乏训练而逐渐萎缩。从长远来看,这不仅影响个体研究者的学术成长,更可能造成学术界的思想同质化——大量AI辅助生成的论文呈现出趋同的表达方式和思维模式,削弱学术研究的多样性和创新活力。

3 应对机制

针对上述三大应用场景中的学术诚信风险,本节从文献检索、数据分析和论文撰写三个维度,提出“技术工具+教育引导+制度约束+能力培养”四位一体的综合治理体系,为医学研究生规范使用生成式AI提供系统性框架。

3.1 文献检索环节的应对机制

针对文献检索环节中由AI幻觉引发的虚假引用与信息干扰风险,首先,在技术工具层面需重塑人机协同范式,将AI从直接提供文献的“检索主体”降级为辅助构建标准检索式的“策略工程师”,并将检索执行权剥离至PubMed等权威数据库,从而在底层逻辑上隔绝文献捏造的可能。其次,在能力培养层面,研究生必须筑牢“人工核查与多源验证”的双重防线,对AI推荐文献的标题、作者及DOI等核心元数据执行严格的源头追溯,确保交叉验证无误后方可纳入引用。同时,在教育引导层面,培养单位应主动破除学生的“技术盲信”,将AI工具伦理纳入学术规范必修课,通过真实案例解剖强化学生对工具边界的认知与批判性思维。最后,在制度约束层面,高校需出台规范指引并划定红线,严禁直接照搬大模型生成的参考文献,强制要求引文具备可溯源的原始检索记录。通过上述四维协同,方能在释放技术红利的时候,系统性捍卫医学科研的可靠性与学术底线。

3.2 数据分析环节的应对机制

在数据分析环节,AI给出的结果不能直接作为研究结论。研究生应保留从数据清理、变量处理到模型选择和假设检验的完整记录,并要求工具说明所采用的方法及其前提。涉及核心结果时,应把代码置于本地环境运行,并用R、Python或其他独立统计软件复算;若复算结果不一致,应回到原始数据和分析流程逐步排查,而不能以模型输出替代验证。这样做的目的,是让每个分

析决定都有依据、每项结果都能够复现。

课程教学和导师指导还需检验学生是否真正理解统计方法。医学统计学课程可结合错误代码、假设条件不满足和偏倚解读等案例,讨论 AI 辅助分析常见的失误。开题和中期考核中,导师应要求学生解释变量设置、方法选择及结果的其他可能含义,避免学生只会运行代码而无法判断分析是否合理。

管理规则应同时说明允许范围和披露要求。AI 可用于代码示例、调试或图表制作,但研究设计、主要统计方法和关键结果解释应由研究者决定。论文方法部分需如实记录所用工具、具体用途、人工修改以及独立验证方式,使读者能够判断分析过程的透明度与可重复性。

3.3 论文撰写环节的应对机制

论文写作中的治理应以作者责任为中心。语法检查、翻译和句式调整可视作语言辅助;研究问题、核心论证、数据解释和结论则不应由 AI 代为完成。学校可使用知网 AIGC 检测或 Turnitin 等系统协助发现需要核查的段落,但检测结果只能作为人工复核的线索,不能单独代替对内容来源和作者贡献的判断。凡经 AI 处理的实质性内容,作者都应回到原始数据和引用文献逐句核对,并说明自己作出的修改。更重要的是,研究生仍需亲自完成论证结构的搭建和主要观点的表达。导师在审阅或预答辩时,可要求学生解释段落之间的推理关系、证据来源及措辞选择,由此判断其是否真正理解并能够为文本负责。培养单位还应把这一要求落实到课程和制度中。学术写作课程除讲授格式外,应增加论证组织、证据评价和改写训练;学位论文则可设置统一的 AI 使用说明,记录工具名称、版本、用途、涉及范围和人工核查过程。对未披露的辅助使用,应先结合文本、过程记录和作者说明进行核实;经认定存在实质性代写或故意隐瞒的,再依学术不端规定处理。

4 医学研究生使用生成式 AI 的行为规范准则

在前述风险识别与应对机制分析的基础上,本节进一步提炼形成医学研究生使用生成式 AI 的行为规范准则,以条文式框架为高校制定相关管理规定、研究生规范自身 AI 使用行为提供具体依据。本准则遵循诚实、透明、负责任三项核心原则,覆盖医学研究生在科研全流程中使用生成式 AI 的主要场景。

准则一(使用边界准则):生成式 AI 仅可作为医学研究生科研工作的辅助工具,不得替代研究者的核心学术训练。具体而言:文献检索阶段,AI 可用于生成检索策略、推荐相关文献,但最终的文献筛选与真实性核查须由研究者独立完成;数据分析阶段,AI 可用于代码生成、统计方

法建议和图表绘制,但研究设计、核心统计方法选择和关键结果解读必须由研究者独立完成;论文撰写阶段,AI 可用于语法检查、句式优化和语言润色,但论文的核心论证、创新点和学术观点必须由研究者独立撰写。各高校应结合学科特点制定具体的 AI 使用边界细则,明确界定“辅助使用”与“替代性使用”的界限。

准则二(披露义务准则):医学研究生在提交任何包含 AI 辅助内容的学术成果时,须以标准化格式履行披露义务。披露内容须包括:所使用的 AI 工具名称及版本号、使用的具体环节与范围、AI 生成内容在文中所占比例,以及对 AI 生成内容的人工核查与修改说明。披露位置建议统一置于学位论文的“致谢”或“附录”部分,期刊投稿须遵循目标期刊的 AI 使用披露政策。未按要求履行披露义务的行为,应视同学术不端予以处理;国内多项高校已出台相关规定,如《复旦大学关于在本科毕业论文(设计)中使用 AI 工具的规定(试行)》明确要求在论文中注明 AI 使用情况。

准则三(核查责任准则):使用者须对 AI 辅助生成的所有内容承担第一核查责任,确保内容的真实性与准确性。对于 AI 推荐的参考文献,须逐一通过权威数据库核实其作者、标题、期刊、卷期页码和 DOI 的真实性;对于 AI 生成的统计结果和数据描述,须通过独立工具进行复现验证;对于 AI 生成的文字内容,须逐句核查其论证逻辑、数据引用和文献标注的准确性。导师须对研究生使用 AI 辅助完成的学术成果履行监督责任,在论文审阅环节将“AI 生成内容核查”纳入审阅要点。

准则四(数据与隐私保护准则):医学研究生在使用生成式 AI 时,须严格遵守数据安全与隐私保护相关规定。禁止将涉及患者隐私的临床数据、未公开的研究数据,以及任何涉密信息上传至公开可用的 AI 平台;使用本地部署或经学校认证的安全 AI 工具开展数据分析时,须确保数据存储与传输的全过程安全。中国科学院科研道德委员会在《关于在科研活动中规范使用人工智能技术的诚信提醒》中明确指出,必须反对滥用 AI 技术危害数据安全、侵犯知识产权、泄露个人隐私等行为^[10]。

准则五(能力培养准则):高校应将 AI 辅助科研工具的使用培训纳入医学研究生培养方案,在提升学生 AI 工具使用能力的同时,确保其核心科研素养不被侵蚀。培训内容应包括:AI 工具的能力边界与局限性、AI 幻觉的识别与防范、AI 辅助内容的核查方法、学术诚信规范与 AI 使用披露要求。高校应在学位论文开题、中期考核和预答辩环节增加对 AI 使用情况的质询,确保研究生在合理使用 AI 工具的同时,其独立科

研能力得到充分训练。

结语

研究显示,生成式AI带来的诚信问题并非彼此孤立。未经核验的引文可能进入综述和讨论,缺少透明记录的分析又可能产生无法复现的结论;如果研究生长期把判断和表达一并交给工具,其写作与科研推理训练也会受到影响。因此,治理重点不宜只放在文本检测上,而应贯穿工具选择、使用记录、结果核验、课程训练和责任追究等环节。基于这些风险,本文从使用边界、披露义务、核查责任、数据保护和能力培养五方面提出行为准则。该框架可为培养单位制定细则提供起点,但其适用效果仍需在不同学科和培养阶段中检验。后续研究可比较各类医学专业的使用情境,评估披露与核查制度的实际执行情况,并通过长期随访观察AI辅助科研对研究生学术能力的影响。检测技术可以作为监督工具之一,其准确性、适用范围及可能产生的误判也应同步评估。

生成式AI可以参与科研过程,却不能承担研究者的学术责任。医学研究生可以利用它减少重复性工作,但对研究问题、方法选择、证据解释和最终表述仍应保持独立判断。只有把可核查、可披露和可追责落实到具体环节,效率收益才不会以牺牲科研诚信为代价。

作者贡献: 张再球与李思杰为本文共同第一作者,对研究工作贡献相当,共同负责研究框架的设计、文献的整理与收集、论文的起草、撰写、修改与编辑工作;文敏负责研究命题的提出与整体研究设计;文敏与何庆南为共同通讯作者,负责论文最终版本的修订、论文整体统筹与协调。

基金项目: 2025年中南大学教育教学改革研究项目(编号:2025jy229)

利益冲突: 本文无利益冲突。

参考文献

- [1] 中华人民共和国教育部. 高等学校人工智能创新行动计划[S]. 北京: 教育部, 2018.
- [2] 陈虹旭, 赖睿, 徐建国, 等. 人工智能赋能神经外科教学: 机遇、挑战与对策[J]. 华西医学, 2024, 39(09): 1470-1473.
- [3] Tozsin A, Ucmak H, Soyuturk S, 等. The Role of Artificial Intelligence in Medical Education: A Systematic Review[J]. Surg Innov, 2024, 31(4): 415-423.
- [4] Combs C D. Emerging Roles of Virtual Patients in the Age of AI[J]. AMA J Ethics, 2019, 21(2): E153-159.
- [5] Nagi F, Salih R, Alzubaidi M, 等. Applications of Artificial Intelligence (AI) in Medical Education: A Scoping Re-

view[J]. Stud Health Technol Inform, 2023, 305: 648-651.

[6] Xu Y, Jiang Z, Ting DSW, 等. Medical education and physician training in the era of artificial intelligence[J]. Singapore Med J, 2024, 65(3): 159-166.

[7] Topaz M, et al. Fabricated citations: an audit across 2.5 million biomedical papers[J]. The Lancet, 2026. DOI: 10.1016/S0140-6736(26)00603-3.

[8] 田贤鹏, 肖智琦. 生成式AI赋能研究生科研写作的学术伦理与风险防控[J]. 现代教育技术, 2024, 34(08): 23-32.

[9] 《生成式人工智能服务管理暂行办法》[J]. 中华人民共和国公安部公报, 2023(05): 2-5.

[10] 中国科学院科研道德委员会. 关于在科研活动中规范使用人工智能技术的诚信提醒[EB/OL]. 中国科学院, 2024-09-10. http://www.cas.ac.cn/sygz/202409/t20240910_5031186.shtml.

[11] Jin Q, Leaman R, Lu Z. PubMed and beyond: biomedical literature search in the age of artificial intelligence[J]. eBioMedicine, 2024, 100: 104988. DOI: 10.1016/j.ebiom.2024.104988.

[12] Alkaiissi H, McFarlane S I. Artificial hallucinations in ChatGPT: implications in scientific writing[J]. Cureus, 2023, 15(2): e35179. DOI: 10.7759/cureus.35179.

[13] Walters W H, Wilder E I. Fabrication and errors in the bibliographic citations generated by ChatGPT[J]. Scientific Reports, 2023, 13: 14045. DOI: 10.1038/s41598-023-41032-5.

[14] Topaz M, et al. Fabricated citations: an audit across 2.5 million biomedical papers[J]. The Lancet, 2026. DOI: 10.1016/S0140-6736(26)00603-3.

[15] Meyer J G, Urbanowicz R J, Martin P C N, et al. ChatGPT and large language models in academia: opportunities and challenges[J]. BioData Mining, 2023, 16: 20. DOI: 10.1186/s13040-023-00339-9.

[16] Sallam M. ChatGPT utility in healthcare education, research, and practice: systematic review on the promising perspectives and valid concerns[J]. Healthcare, 2023, 11(6): 887. DOI: 10.3390/healthcare11060887.

[17] Bond M, Khosravi H, De Laat M, et al. A meta systematic review of artificial intelligence in higher education: a call for increased ethics, collaboration, and rigour[J]. International Journal of Educational Technology in Higher Education, 2024, 21: 4. DOI: 10.1186/s41239-023-00436-z.

[18] 王少. ChatGPT与学术不端治理: 挑战与应对[J]. 科技进步与对策, 2023, 40(23): 103-110. DOI:

10.6049/kjbydc.2023020435.

[19] 杨顺. ChatGPT 等生成式人工智能对学术诚信的挑战及应对 [J/OL]. 上海市法学会东方法学, 2024. https://www.thepaper.cn/newsDetail_forward_28800629.

[20] 彭甲超, 田嘉莉, 杨爽, 等. 大型语言模型对学术型研究生学术规范和论文写作的影响及挑战——以 ChatGPT 为例 [J]. 产教融合, 2024, 6(4): 11-30. DOI: 10.6938/iie.060402.

[21] Dwivedi Y K, Kshetri N, Hughes L, et al. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy[J]. International Journal of Information Management, 2023, 71: 102642. DOI: 10.1016/j.ijinfomgt.2023.102642.

[22] Cheng A, Calhoun A, Reedy G. Artificial intelligence-assisted academic writing: recommendations for ethical use[J]. Advances in Simulation, 2025, 10: 22. DOI:

10.1186/s41077-025-00350-6.

[23] Cotton D, Cotton P A, Shipway J R. Chatting and cheating: ensuring academic integrity in the era of ChatGPT[J]. Innovations in Education and Teaching International, 2023, 61(2): 228-239. DOI: 10.1080/14703297.2023.2190148.

[24] Chan C K Y, Hu W. Students' voices on generative AI: perceptions, benefits, and challenges in higher education[J]. International Journal of Educational Technology in Higher Education, 2023, 20: 43. DOI: 10.1186/s41239-023-00411-8.

[25] Taloni A, Scoria V, Giannaccare G. Large Language Model Advanced Data Analysis Abuse to Create a Fake Data Set in Medical Research. JAMA Ophthalmol. 2023 Dec 1;141(12):1174-1175. doi: 10.1001/jamaophthalmol.2023.5162. PMID: 37943569; PMCID: PMC10636646.