

医疗机构部署医学人工智能技术的评估考量

吉萍¹ 祝丹娜¹ 肖平¹ 梁真²

¹深圳北京大学香港科技大学医学中心临床研究所, 深圳 518036; ²北京大学深圳医院, 深圳 518036

通信作者: 梁真, Email: liangzhen@pkusz.com, 电话: 0755-83923333

【摘要】目的 本文围绕医疗机构如何开展医学 AI 部署应用进行评估, 期望能对 AI 技术在我国医学领域安全、可靠及有价值的应用有所启示。**方法** 梳理国内外医学 AI 法规规范中评估相关要求, 以及机构开展医学 AI 应用评估的实践模式。**结果** 医疗机构在部署 AI 应用时, 建议应设立 AI 管理部门, 组建跨学科多元化评估团队, 构建多维度评估指标体系, 实施分阶段评估。**结论** 医疗机构要关注 AI 系统“安全有效性”“部署必要性”及“临床有用性”等, 确保医学 AI 技术安全可靠、符合伦理, 实现提质增效的目标。

【关键词】 医学人工智能; 医疗机构; 评估; 伦理

基金项目: 深圳市“医疗卫生三名工程”项目资助(SZSM202211042)

【中图分类号】 R197; R-05 **【文献标识码】** A DOI:10.3760/cma.j.cn113565-20251010-00268

Evaluation considerations for the deployment of medical artificial intelligence technologies in medical institutions

Ji Ping¹, Zhu Danna¹, Xiao Ping¹, Liang Zhen²

¹Clinical Research Institute, Shenzhen Peking University-Hong Kong University of Science and Technology Medical Center, Shenzhen 518036, China; ²Peking University Shenzhen Hospital, Shenzhen 518036, China

Corresponding author: Liang Zhen, Email: liangzhen@pkusz.com, Tell: 0086-755-83923333

【Abstract】Objective This study focused on the deployment evaluation of medical AI in healthcare institutions, aiming to provide insights for the safe, reliable, and valuable application of AI technology in healthcare field. **Methods** The research systematically reviewed the evaluation requirements in domestic and international medical AI regulations, the characteristics of AI technology, the adopted evaluation models by institutions. **Results** When deploying AI applications, dedicated AI management departments, interdisciplinary evaluation teams, comprehensive assessment criteria and evaluations should be setup. **Conclusions** Healthcare institutions should pay attention to AI systems' safety and effectiveness, deployment necessity, and clinical utility. This ensures the safety, reliability, and ethical compliance of medical AI technologies, ultimately achieving the goal of improving quality and efficiency.

【Key words】 Medical artificial intelligence; Medical institutions; Evaluation; Ethics

Fund program: Supported by Shenzhen's Sanming Project-Projects(SZSM202211042)

DOI:10.3760/cma.j.cn113565-20251010-00268

中国一直在大力推动人工智能(Artificial Intelligence, AI)产业的发展, 医疗健康领域成为 AI 技术落地的核心场景之一。当前, 医疗数字化智能化转型加速推进的背景下, 医学 AI 技术的应用潜力正呈现爆发式增长。2024 年 11 月, 国家卫生健康委、国家中医药局、国家疾控局联合发布了《卫生健康行业人工智能应用场景参考指引》, 系统规划了卫生健康领域人工智能应用方向, 涵盖“人工智能+”医疗服务管理的 15 种具体应用场景, 包括医学影像智能辅助诊断、智能辅助治疗、智能辅助决策、智能预问诊及智能辅助质控与管理等。随着促进应用的政策引导进一步加速了医疗 AI 的落地进程。医疗机构在引入并部署 AI 系统的过程中, 首要关

注部署的 AI 系统是否安全有效。同时, 医疗机构需确保部署决策符合医疗服务的核心需求, “从临床需求导向”出发, 分析 AI 系统能否解决现有诊疗流程中的难点堵点问题。然而, AI 技术作为新兴技术领域, 医疗机构内精通 AI 知识的专家相对匮乏, 相关方管理经验少。医疗机构追逐技术热点, 可能引入尚未得到充分验证的 AI 系统, 或未结合机构自身诊疗需求, 未能有效融入临床路径等, 导致“为 AI 而 AI”的资源浪费、误诊和不当诊断等医疗风险。

笔者团队参与调研医疗机构 AI 应用场景工作时发现, 虽然医疗机构积极推动部署各类模型, 快速推进 AI 产品的研发与应用。然而, AI 部署出现“落地不畅”。一方面, AI 的评估主要是在实验室环

境中评估模型的准确性和稳健性。从“实验室”研发构建的 AI 模型部署到复杂多变的“临床真实”应用场景中时,随着临床实践的变化、真实患者特征与训练样本数据的差异、AI 工具嵌入临床工作流程的程度、医护人员对 AI 工具的接受度、其他不确定情境因素等,都可能会影响 AI 工具在医疗机构的有效、安全、公平和公正应用^[1-3]。另一方面,医疗机构对 AI 技术的监管往往涉及多个部门,包括医务处、设备管理部门、信息部、病案室、伦理委员会及相关临床科室等,各主体权责边界模糊、缺乏协作机制,内部 AI 管理机制不清晰,导致 AI 引入评估与后期管理陷入“多头负责却无人统筹”的困境。

医疗机构应理性全面评估 AI 的应用部署,围绕“部署的 AI 系统是否安全有效”“新 AI 系统应不应该被部署”及“如何将 AI 技术与临床诊疗整合,实现提质增效”等问题。因此,迫切需要医疗机构探索建立 AI 技术多维度、分阶段、多元化多角度评估框架,制定明确、可操作的评估标准和流程,确保 AI 技术在医疗领域的安全、可靠和有效应用。2025 年 11 月,国家卫健委等五部门印发《关于促进和规范“人工智能+医疗卫生”应用发展的实施意见》指出,要制定评估标准,开展应用监测评估。

1 医学 AI 法规与规范中的评估要求

全球各国政府与国际组织、学术界和工业界共同积极寻求 AI 治理的框架、指南和原则。自 2021 年,WHO 已发布一系列医学 AI 伦理治理相关指南中都强调“医疗 AI 技术在部署前,需通过严格的安全性、准确性及有效性评估,并建立持续监测机制以识别潜在风险;从模型设计、训练、内部验证到外部验证的完整开发流程进行相应评估;强调由监管机构评估和批准用于医疗领域的多模态大模型和应用程序,应建立医疗机构应对技术失误或滥用导致的损害责任机制^[4-7]”。欧盟发布 AI 法规及报告中建议医学 AI 技术属于高风险,所需的测试和监管就越严格,需要更全面、多维度、分阶段的评估方法,除准确性外,还应评估公平性、可用性、可解释性、透明性和可行性等^[8]。

2021 年,国家药监局发布了《人工智能医用软件产品分类界定指导原则》和《人工智能医疗器械注册审查指导原则》规范 AI 医疗器械的技术审评要求,针对不同算法类型(如深度学习、机器学习)及其组合应用,需开展整体验证与确认工作,特别关注深度学习等黑盒算法的透明度提升及新算法(如小样本学习、弱监督学习)的专项风险评估,以确保算法在预期用途和使用场景下的安全有效。AI 医疗器械的软件安全性级别应与其潜在风险相匹配,并通过综合判定预期用途(如用途类型和临床紧迫性)、使用场景(如适用人群和疾病严重程度)及核心功能(如算法类型和输入输出接口)来确定风险等级^[9-10]。中国国家卫生健康发展研究中心已开展了 AI 医疗技术临床应用评估指南的研究与应用,形成了包含技术特性、安全性、经济性等维度的评估指标体系及评估流程,为医保定价、技术监管及临床合理应用提供了循证依据^[11]。

2024 年,深圳市卫生健康委员会印发《关于加强“人工智能+医疗健康”应用安全管理工作的通知》,明确“谁使用谁负责”原则,要求医疗卫生机构建立 AI 应用管理制度,覆盖责任归属、追究及风险管控机制。使用单位需在应用前开展论证评估,强化过程管理与应用监测;AI 使用单位应设立专门部门统筹“AI+医疗健康”产品和服务管理,需详细记录产品注册信息、技术安全性、算法原理、应用合法性、场景适配性、预期效果及临床团队资质等内容,确保全流程可追溯、可监管。

2 机构开展医学 AI 应用评估的实践探索

国际组织、政府、企业等均积极从 AI 全生命周期视角,对医学 AI 技术的部署进行全方位测试与评估,来确保负责任地使用 AI。各国政府主要通过设立标准化评估平台、评估框架或采用沙盒测试等进行监管。传统测试和评估策略主要包括侧重于模型属性的报告指南和检查清单。在医学机构层面,美国杜克大学、斯坦福医学院、新加坡保健集团和中国台湾高雄医科大学附属中和纪念医院等已建立开展医学 AI 应用评估的机制,具体内容详见表 1^[12-14]。

表 1 不同机构 AI 评估模式比较

对比维度	杜克大学	斯坦福医学院	新加坡保健集团	中国台湾高雄医科大学附属中和纪念医院
体系/制度名称	基于算法的临床决策支持 (Algorithm Based Clinical Decision Support: ABCDS) 监督体系	验证和测评 AI 工具的评估模型 (Fair, Useful, and Reliable AI Models, FURM)	AI 治理制度	AI 模型临床部署应用管理办法

续表 1

对比维度	杜克大学	斯坦福医学院	新加坡保健集团	中国台湾高雄医科大学附属中和纪念医院
目标	为临床诊疗和运营中部署 AI 提供评估和监测需求	“技术—伦理—经济”多维评估验证 AI 工具,为引进 AI 技术的决策提供依据	遵循新加坡《医疗人工智能指南》	促进 AI 在临床场景的合规应用
评估原则或维度	透明性和问责制、临床价值和安全性、公平性、可用性和采纳度、合规性等 5 项伦理原则	技术有用性、7 个伦理原则(责任、公平、可追溯性、可靠性、治理、不伤害、自主性)、财务指标	基于风险的分级分类管理	技术可行性、资源配置需求、临床适配性等多维度评估,关注效能、法规合规等
评估方式	从模型开发、静默评估、有效性验证到持续监控的全过程“反馈—改进”方式	分“是什么与为什么”“怎么做”“影响”3 个阶段开展评估验证,各阶段有对应不同的具体评估任务	机构内多元化部门参与,设立 AI 办公室,有 AI 评估委员会与 AI 治理监管委员会,部署高风险 AI 系统前多学科伦理委员会审查	设立 AI 临床部署审查小组,分初始审查、自评风险、多维度技术审查、部署与运行管理、模型监控与反馈等阶段
评估人员构成	临床专家、医院管理者、生物统计学家、AI 工程师、数据科学家、法学和医疗公平方面	开发者、患者代表、数据科学团队、临床服务团队、机构领导层	临床专家、医院管理者、AI 工程师、数据科学家等	AI 办公室、审查小组、信息部门、相关领域专家、临床使用科室

3 建议

医疗机构引入 AI 技术是为了提升医疗服务质量与效率,要关注 AI 系统“安全有效性”“部署必要性”及“临床整合路径有效性”等,构建“多维度、分阶段、多元化”的系统性评估框架,遏制不成熟产品的滥用,规避技术风险,实现提质增效的目标。

3.1 建立 AI 管理部门

医疗机构应设立独立的 AI 专项管理办公室,负责统筹促进医学 AI 在本机构的规范应用,包括协调临床科室、信息部门和外部供应商等,完善多部门协作机制,促进开发者、使用者与评估者的沟通与交流。通过建立分级审批机制,评估流程与指标并完善管理制度,涵盖项目申报、审批、实施和实时监测 AI 系统的运行效果。组织对 AI 产品的技术可行性和可靠性进行评估,确保 AI 应用符合相关规范并实现预期目标。同时,负责组织的 AI 知识培训和科普宣传工作,提升医务人员和患者对医疗 AI 的客观认知,AI 应用能力和风险意识。

3.2 组建跨学科多元化评估团队

组建由多学科专业人员构成的 AI 临床部署审查评估团队。团队成员构成应覆盖:一是临床专家,负责评估 AI 产品与临床诊疗流程的适配性,如是否符合医师使用习惯、是否能提升诊疗效率、是否增加患者负担、诊疗价值和应用风险。二是 AI 技术

专家(算法工程师、数据科学家),从技术层面审核数据安全防护(加密技术、访问权限管理)、算法稳定性、泛化能力及可解释性等。三是医学伦理和法学,聚焦数据隐私保护、知情同意机制、特殊人群保护、应用边界、伦理合规性和法律责任界定等风险应对。四是医院管理与成本核算专家,开展成本效益分析和资源配置合理性,评估 AI 产品的投入产出比。

3.3 构建多维度评估指标体系

形成“技术安全可靠—伦理合规—成本效益合理”,涵盖系统性能、临床效果和价值的多维度评估体系,确保 AI 应用安全、合规且符合临床需求。(1)技术评估:首要关注部署的 AI 系统是否安全有效。评估潜在的技术有用性、技术成熟度、技术性能与可靠性的核心指标,提供决策依据的可视化说明,通过测试验证 AI 的敏感性、特异性及准确率等关键指标,确保在临床场景中稳定运行;检查与现有医疗设备、电子系统的数据接口兼容性。评估部署 AI 模型融入工作流程所需的改变、IT 基础设施和人员配置是否充分,确保其能够融入临床工作流程。(2)伦理方面:基于通行医学伦理准则,系统地判断是否会引发伦理问题?披露与伦理相关的注意事项和措施,以帮助用户建立信任。针对已发现问题是否有降低伦理风险的措施?实施过程评估或引入定性研究作为重要的补充研究类型,重视理解医生和患者对医学 AI 的态度、

认识、观点、期望、体验和选择考虑因素,医学 AI 对医疗服务的影响。(3)成本效益评估:需确保部署决策符合医疗服务的核心需求,“从临床需求导向”出发,分析 AI 系统能否提升现有诊疗流程中的质量或效率。开展全面的成本效益分析,涵盖采购成本、部署成本以及运维成本费用,与传统诊疗方案对比,评估 AI 应用的长期机构运行、人均诊疗成本变化等。

3.4 实施分阶段评估流程

建立分阶段、全周期审查流程:(1)申请与受理阶段:由 AI 办公室负责受理 AI 应用项目申请,对申请者申请材料的完整性和合规性进行初步审核,确认项目符合申请受理范围及风险等级。(2)引入前的综合评估阶段:由多学科审查小组从技术可行性、可靠性、成熟度、资源配置需求、与现有医疗系统的兼容、对所需的资源投入和成本效益进行综合评估。(3)部署与监控阶段:申请获批后,在特定科室或场景开展小范围试点,跟踪 AI 产品的实际应用效果,由 AI 管理办公室联合信息部门和临床使用科室共同评估 AI 应用的实际效果,分析 AI 系统对诊疗质量和效率的实际影响;开展用户满意度调查和定性研究,了解医务人员和患者对 AI 应用的接受程度和改进建议,并定期监测汇总评估结果,由 AI 办公室会同相关领域专家确认产品的稳定性与可靠性。关键在于打破“技术与临床脱节”的壁垒,构建“人机协同”的诊疗模式,需建立“用户反馈—技术迭代”的闭环机制,根据医师使用体验优化 AI 系统的操作界面、交互逻辑,提升临床适用性。一旦出现超过监测标准值或可接受风险的情形,将立即启动停用及重新调校程序至改善完成。通过实时监测系统的运行数据,评估其稳定性和可靠性;收集临床应用数据,分析 AI 系统对诊疗质量和效率的实际影响;开展用户满意度调查,了解医务人员和患者对 AI 应用的接受程度和改进建议。如果 AI 部署不符合医师使用习惯、不匹配医院信息系统架构,可能反而增加临床工作负担。构建涵盖系统性能、临床效果、用户反馈的多维监控体系。通过实时监测系统的运行数据,评估其稳定性和可靠性;收集临床应用数据。从医疗质量安全、个人隐私和数据安全等方面开展穿透式监管,加强动态监测和预警。

4 小结

在促进医学 AI 技术应用过程中,医疗机构需基于医学 AI 技术特点与规律,对医学 AI 产品应用进行全面评估,通过设立 AI 管理部门,组建跨学科

多元化评估团队,构建多维度评估指标体系,并实施分阶段评估,确保医学 AI 技术在医疗机构安全可靠、符合伦理的应用。只有在技术可靠性与临床适配性之间实现动态平衡,让医学 AI 真正摆脱概念炒作的泡沫风险,实现从“技术创新”到“价值创造”的根本转变,真正成为提升医疗服务质量、保障人民健康的工具。

利益冲突 所有作者均声明不存在利益冲突

作者贡献声明 吉萍:研究设计、文章撰写;祝丹娜:查阅文献,参加讨论;肖平:研究指导、论文审阅;梁真:研究指导、论文审阅

志谢 感谢新加坡保健集团伦理办公室伍德全博士提供的信息与指导

参 考 文 献

- [1] British Medical Association. Principles for Artificial Intelligence (AI) and its application in healthcare [EB/OL]. (2024-10-01) [2025-06-06]. <https://www.bma.org.uk/media/njgfbmnn/bma-principles-for-artificial-intelligence-ai-and-its-application-in-healthcare.pdf>.
- [2] Alex John London. Artificial intelligence in medicine: Overcoming or recapitulating structural challenges to improving patient care? [J]. Cell Reports Medicine, 2022, 3(5):100622. DOI:10.1016/j.xcrm.2022.100622.
- [3] Eric Wu, Kevin Wu, Roxana Daneshjou. How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals [J]. Nature medicine, 2021, 4(1). DOI:10.1038/s41591-021-01312-x.
- [4] World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance [EB/OL]. (2021-06-28) [2024-10-06]. <https://www.who.int/publications/i/item/9789240029200>.
- [5] World Health Organization. Generating Evidence for Artificial Intelligence Based Medical Devices: A Framework for Training Validation and Evaluation [EB/OL]. (2021-11-17) [2024-10-06]. <https://www.who.int/publications/i/item/9789240038462>.
- [6] World Health Organization. Regulatory considerations on artificial intelligence for health [EB/OL]. (2023-10-19) [2024-10-06]. <https://iris.who.int/bitstream/handle/10665/373421/9789240078871-eng.pdf>.
- [7] World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models [EB/OL]. (2024-01-18) [2024-10-06]. https://iris.who.int/handle/10665/375579?search-result=true&query=artificial+intelligence+for+health&.scope=&.rpp=10&.sort_by=score&.order=desc.
- [8] European Parliamentary Research Service (EPRS). Artificial intelligence in healthcare: Applications, risks, and ethical and

- societal impacts [EB/OL]. (2022-06-01) [2025-06-12]. [https://www.europarl.europa.eu/stoa/en/document/EPRS_STU\(2022\)729512](https://www.europarl.europa.eu/stoa/en/document/EPRS_STU(2022)729512).
- [9] 国家药监局. 人工智能医用软件产品分类界定指导原则 [EB/OL]. (2021-07-01) [2025-06-06]. <https://www.nmpa.gov.cn/xxgk/ggtg/ylxqggtg/ylqxqggtg/20210708111147171.html>.
- [10] 国家药监局器审中心. 人工智能医疗器械注册审查指导原则 [EB/OL]. (2022-03-07) [2025-06-06]. <https://www.cmde.org.cn/xwdt/shpgzgg/gztg/20220309090800158.html>.
- [11] 刘克军,肖月,邱英鹏,等. 我国人工智能医疗技术临床应用评估指南研究与应用 [J]. 医学信息学杂志, 2023, 44(10): 16-21.
- [12] Nicoleta J Economou-Zavlanos, Sophia Bessias, Michael P Cary, et al. Translating ethical and quality principles for the effective, safe and fair development, deployment and use of artificial intelligence technologies in healthcare [J]. Journal of the American Medical Informatics Association 2024, 31(3): 705-713. DOI: 10.1093/jamia/ocad221.
- [13] Callahan A, Steinberg E, Fries J, et al. Standing on FURM ground - A framework for evaluating Fair, Useful, and Reliable AI Models in healthcare systems [J]. E-Print arXiv, 2024(10): 687-691. DOI: 10.48550/arXiv.2403.07911.
- [14] The Ministry of Health (MOH), the Health Sciences Authority (HSA) & Integrated Health Information Systems (IHIS). Singapore publishes Artificial Intelligence in Healthcare Guidelines (AIHGIE) [EB/OL]. (2021-11-03) [2025-06-20]. <https://andamanmed.com/singapore-publishes-artificial-intelligence-in-healthcare-guidelines-aihgrie/>.
- [15] Zeng D, Qin Y, Sheng B, et al. DeepSeek's "Low-Cost" Adoption Across China's Hospital Systems: Too Fast, Too Soon? [J]. Journal of the American Medical Association, 2025, 333(21). DOI: 10.1001/jama.2025.6571.

(收稿日期:2025-10-10)

关于不法分子冒充《中华医学科研管理杂志》编辑部工作人员的声明

近期,有不法分子冒充中华医学会系列杂志多家期刊(或本刊杂志名)编辑部的工作人员,向已发表论文的通信作者、作者发送诈骗邮件、短信等,要求添加微信好友,其理由主要包括但不限于以下内容:

- (1) 谎称已刊发稿件需要核查作者信息;
- (2) 谎称已刊发稿件需要进行文章抽查、数据抽查;
- (3) 谎称已刊发稿件因数据库调整需要更换链接;
- (4) 谎称近期已刊发稿件需要确认刊期;
- (5) 邀请专家加入学术群。

在此,本刊编辑部郑重声明:

1. 本刊对外发送邮件的工作邮箱是 kgzz@bjmu.edu.cn / cjmsrm@cmaph.org。
2. 本刊工作人员不会以匿名形式与作者联系,更不会私下要求作者缴纳任何费用。任何以编辑部名义要求向个人账户转账的行为均属于电信诈骗。

请广大作者提高警惕,注意甄别,如收到可疑邮件、短信等,请第一时间联系本刊编辑部进行核实,以免上当受骗,造成损失。本刊编辑部联系方式:

杂志官网: <https://zhyxkyglzz.yiigle.com/>

编辑部地址、邮编:北京市海淀区学院路 38 号,100191

联系方式:冯编辑、谢编辑,010-82802696;李编辑,010-82802217

《中华医学科研管理杂志》编辑部

2025 年 10 月