

我国人工智能医疗技术临床应用评估指南 研究与应用*

刘克军 肖月 邱英鹏 史黎炜 田雪晴 游茂 付强

(国家卫生健康委卫生发展研究中心 北京 100044)

〔摘要〕 **目的/意义** 建立我国人工智能医疗技术临床应用评估管理工作体系, 支持临床合理使用决策。**方法/过程** 采用文献分析、专家咨询和实证研究方法, 总结实际应用经验, 形成人工智能医疗技术临床应用评估指南文本。**结果/结论** 研究探索人工智能医疗技术定义与分类、主题遴选、评估方法选择、指标体系构建、质量控制要求, 形成我国人工智能医疗技术临床应用评估指南。经实证研究, 该评估指南有助于支持临床合理使用决策, 为人工智能医疗技术临床应用评估工作规范开展奠定坚实基础。

〔关键词〕 人工智能; 医疗技术; 指南

〔中图分类号〕 R-058 **〔文献标识码〕** A **〔DOI〕** 10.3969/j.issn.1673-6036.2023.10.003

Study and Application of the Clinical Application Evaluation Guideline for Artificial Intelligence Medical Technology in China

LIU Kejun, XIAO Yue, QIU Yingpeng, SHI Liwei, TIAN Xueqing, YOU Mao, FU Qiang

China National Health Development Research Center, Beijing 100044, China

〔Abstract〕 **Purpose/Significance** To establish the clinical application evaluation and management system of artificial intelligence (AI) medical technology in China, and to support clinical rational use decision-making. **Method/Process** Literature analysis, expert consultation and empirical research methods are used to summarize practical application experience and form a guide text for clinical application evaluation of AI medical technology. **Result/Conclusion** The paper studies and explores the definition and classification of AI medical technology, topic selection, evaluation method selection, index system construction, quality control requirements, and forms a guideline for clinical application evaluation of AI medical technology in China. Through empirical research, the evaluation guideline helps to support clinical rational use decision-making, and lays a solid foundation for the standard development of clinical application evaluation of AI medical technology.

〔Keywords〕 artificial intelligence (AI); medical technology; guideline

1 引言

人工智能 (artificial intelligence, AI) 是推动人类进入智能时代的决定性力量, 正在引领新一轮产业

变革, 加快发展新一代人工智能是事关我国能否抓住新一轮科技革命和产业变革机遇的战略问题。我国于 2017 年颁布《新一代人工智能发展规划》^[1], 2022 年印发《“十四五”数字经济发展规划》^[2], 2023 年印发《数字中国建设整体布局规划》^[3], 推动引导 AI 在多个领域迅速发展。在健康中国建设进程中, AI 在卫生健康领域获得广泛应用, 显示出提升效率与质量、赋能基层医疗服务的重要作用, 但同时也面临应用管理的新挑战^[4-5]。

〔修回日期〕 2023-10-23

〔作者简介〕 刘克军, 研究员; 通信作者: 付强, 研究员。

〔基金项目〕 科技创新 2030——“新一代人工智能”重大项目 (项目编号: 2020AAA0105002)。

当前我国 AI 临床应用监管主要面临 3 方面挑战。一是 AI 的临床推广应用仍面临漫长适应过程,待建立相应的使用规范和质控标准助力高价值产品在临床有效应用。二是 AI 技术本身监管难题尚未破解,临床应用风险应强化管控。三是医疗机构在使用 AI 医疗技术过程中面临产品性能不稳定、数据安全和患者隐私保护难度大、产品价格昂贵等多重问题,应建立统一评估标准帮助遴选适宜的产品和技术。针对上述问题,国家卫生健康委卫生发展研究中心牵头开展 AI 临床应用评估指南开发研究,结合科学卫生治理要求,形成统一评估标准及方法工具,助力 AI 临床应用管理决策。

2 指南开发步骤及方法

2.1 基础研究

2021 年 2—5 月,利用文献研究、案例分析和专家咨询等方法,在总结 AI 医疗技术临床应用需求现状、主要问题及国内外评估研究实践经验基础上,建立 AI 临床应用评估指南基本框架,主要包括 AI 医疗技术定义与分类、主题遴选与决策场景、评估设计与实施技术标准及质量控制方法 3 部分。

2.2 指南开发研究

按照科学技术部重大课题要求,评估中心组织 11 家医疗机构及科研院所等项目合作单位专家力量,深入研究并开发指南文本。过程中广泛征求管理、临床、医工、医学伦理、循证医学、卫生技术评估、AI 医疗技术研发等领域权威专家意见。2021 年 7 月—2023 年 6 月重点围绕以下问题展开研究。一是 AI 医疗技术定义与分类研究。基于文献研究和专家咨询,结合卫生、药品监管行政主管部门发布的相关政策文件,确定 AI 临床应用的多种场景,对以 AI 为载体的医疗技术进行定义,并基于“应用场景”“临床风险”等原则进行分类。二是优先评估 AI 选择方法(即评估主题遴选)。基于文献研究及临床应用核心问题分析,探索 AI 医疗技术主题遴选方法及指标体系,确定了包括重要性、相关性和可评估性 3 个维度的主题遴选决策工具;结合

专家咨询意见,筛选确定我国 AI 医疗技术潜在应用场景。三是评估设计与实施技术标准。基于国内外通用的卫生技术评估(health technology assessment, HTA)方法框架^[6-8],确定 AI 医疗技术临床应用评估设计实施技术标准和要求,提出技术特性、创新性、伦理公平性、临床价值(安全性、有效性、效率)、经济性等主要维度指标推荐意见及测量方法,并提出核心数据收集及质控要求。

2.3 指南验证研究

基于初步形成的指南文本,于 2021 年 7 月—2022 年 12 月选择肿瘤、呼吸系统、内分泌等重大疾病救治场景,联合北京友谊医院、安徽省立医院、四川大学华西医院等试点机构,以疾病筛查、辅助诊断、合理用药管理等为具体试点场景,分批次选择试点进行实证研究,探索特定领域临床应用评估的具体方法工具要求,验证和完善了评估指南的适用性和科学性,为完善指南文本提供依据。目前已与北京友谊医院联合完成了“冠状动脉 CT 血管成像人工智能辅助诊断技术的临床应用评估”,有关评估结果支持 AI 应用定价报销决策,经北京市科学技术委员会认可后推动后继医联体应用示范研究。实证研究表明指南对规范评估过程和确保评估结果可用性有重要作用,可助力临床服务和管理循证决策机制的构建。

2.4 指南修订完善

2022 年 5—10 月,经过 10 轮大范围专家咨询,结合试点评估验证结果反馈,课题组修订指南文本及相关工具,形成 AI 医疗技术临床应用评估指南初稿。

3 主要发现

3.1 AI 医疗技术定义与分类

目前国内缺乏对 AI 医疗技术的统一定义,导致实际监管工作难以有效开展。原国家卫生健康委医政医管局发布的限制类医疗技术清单中仅包括 AI 辅助治疗技术,涵盖范围有限。首先,纳入监管的 AI 辅助治疗技术专指机器人手术辅助系统相关技

术,但目前手术机器人并不具备 AI 功能,而人工智能的手术导航产品、放疗靶区勾画产品等其他 AI 辅助治疗技术并未纳入技术监管范围。其次,纳入监管的人工智能辅助诊断技术是指基于计算机辅助诊断软件及临床决策支持系统的独立软件,并不包括人工智能嵌入式临床诊断与治疗仪器设备。最后,目前部分 AI 非医疗器械(如知识图谱)也在临床上使用,但是并未纳入技术监管范围。

根据本次指南开发的目的,参考我国各相关决策部门对医疗器械、医疗技术、人工智能辅助诊疗技术等的定义^[9-14],应从临床应用评估角度出发,借鉴国家卫生健康委员会《医疗技术临床应用管理办法》^[8]中的医疗技术定义并适度扩展,将 AI 医疗

技术定义为以健康维护促进及相关决策支持为目的,以人工智能技术为支撑的医疗技术。同时,确定以预期用途、使用场景、潜在风险为标准的分类分级原则。其中,一级分类为 AI 辅助诊断和 AI 辅助治疗技术;二级分类为导航规划类、图像识别类、决策支持类和手术类技术。上述各类技术再根据潜在风险和获益分为高、中、低 3 级。

3.2 评估流程

指南在卫生技术评估通用流程^[8]基础上,结合 AI 医疗技术临床应用特点,形成 8 步评估流程,见表 1。研究者可结合实际评估类型和角度、评估承担方条件及证据可得性情况酌情进行有关步骤调整。

表 1 AI 医疗技术临床应用评估基本步骤

步骤	主要研究活动
需求分析	开展文献研究及相关利益者分析,确定评估需求
明确问题	按照 PICOS 原则科学设计研究问题
评估设计	撰写研究计划,包括定义背景、研究问题、内容方法、活动内容
原始研究	通过机构调查等进行 AI 医疗技术临床应用数据收集分析;咨询相关专家,获得补充证据
二次研究	按照系统评估一般性要求,系统总结证据并进行质量评估;鉴于 AI 医疗技术相关文献较少,在咨询方法学和临床专家基础上酌情降低证据等级要求
证据汇总	对文献证据、专家咨询结果、一手数据分析结果等进行汇总分析;证据有争议或矛盾时可咨询有关专家意见
证据呈现	撰写技术报告及针对决策者的摘要报告
论证评审	组织专家针对评估方法及结果进行论证,判断证据转化可行性

注:PICOS 来自多个研究要素的缩写,包括研究对象(patient/population)、干预措施(intervention)、比较因素(comparison/control)、结局(outcome)、研究设计方案(study design)。

3.3 评估设计要点

3.3.1 评估主题遴选 围绕决策问题分析待评 AI 医疗技术主要功能和临床需求情况,就待解决问题的重要性、决策相关性和证据本身完备性、可评估性,广泛征求政策、临床、技术专家意见,最后确定待评估的问题。

3.3.2 确定决策场景 明确 AI 医疗技术应用场景,建议结合特定应用场景及委托方具体决策目的确定评估场景。主要场景可分为以下 5 类。一是国家卫生健康部门 AI 医疗技术监管政策涉及限制或鼓励类清单调整及更新,此类问题优先关注技术风

险和使用安全性、临床获益(疗效和效率)及技术操作要求。二是针对适宜技术遴选推广决策,国家卫生健康部门需要考虑 AI 医疗技术的可靠性(安全有效)和适宜性,研究医患双方接受度和推广配套条件。三是针对医保定价及报销决策,AI 医疗技术能否纳入医保基金支付范围,需要考虑开展基于卫生经济学评估研究,测算成本、评估成本效果,并分析其对医保预算产生的影响。四是针对 AI 医疗技术上市后质量监管决策,国家药品监督管理部门需要对产品临床表现进行长期跟踪评估,此类场景侧重于产品在实际临床使用中的质量稳定性、远期安全性、在更大人群范围内的有效性、适应证及

适用人群的扩展或限制、临床需求研判。五是针对公立医院内控运营决策,医疗卫生机构要考虑 AI 医疗技术的引入对学科发展的影响、临床获益及运维成本。

3.3.3 评估方法选择 AI 医疗技术临床应用评估建议采用 HTA 作为基本方法。按照国内外通用方法学指南^[8],卫生技术评估目前基本可分为完整卫生技术评估 (full HTA)、快速卫生技术评估 (rapid HTA) 及医院卫生技术评估 (HB-HTA) 3 类。当已有文献研究不能满足评估需求时,在时间和经费投入均有保障的情况下,推荐采用完整卫生技术评估,结合原始研究进行评估分析。开展原始研究建议首选基于医院日常工作临床数据的真实世界研究。在真实世界研究中,当数据可以满足决策需求时,建议采用回顾性观察研究设计;当回顾性研究所采集数据存在各种偏倚,数据质量难以保证,已知或已测、未知或不可测量的混杂因素较难识别时,建议采用前瞻性研究设计。真实世界研究设计方法应予以考虑,研究者也可选择其他临床或流行病学研究设计方法。

3.3.4 指标体系构建 在 AI 医疗技术指标定义方面,应根据 AI 医疗技术临床使用特点,包括 AI 医疗技术类型、产品性能、预期用途、应用时的配套要求等,并按照相关性、可靠性、可用性等原则,同时兼顾临床数据可得性情况,选择核心评估指标。不同 AI 医疗技术的临床使用特点不同,指标选择的侧重点也不同。(1) 技术特性指标。评估重点是明确技术类型、产品核心性能、应用要求以及临床需求,以判断技术在临床中的定位、性能优势、操作难度以及应用范围。(2) 创新性指标。建议从临床创新性、服务创新性和产业创新性 3 个维度进行评估。如 AI 医疗技术是否满足其他临床技术未能满足的健康需求、是否能提高区域内诊疗服务资源利用效率、是否为自主研发或全球首创等。(3) 伦理公平性指标。推荐参考权威组织(世界卫生组织、英国国家卫生与临床优化研究所等)的伦理评估分析框架,通常包括保护人类自主权、促进人类福祉和安全以及公共利益、确保透明度/可解释性/可理解性、

培养责任感和问责制、确保包容性和公平性、提高响应性和可持续性 6 方面^[15]。(4) 安全性指标。应重点参考待评 AI 医疗产品说明书中的安全提示信息,以及中国国家药品监督管理局发布的公告、主要国家/地区药监部门发布的撤市/警告信息、AI 医疗产品生产企业产品召回信息,确定具体评估指标,如性能稳定性、诊断符合率、并发症发生水平等。(5) 有效性指标。依据医药卫生领域行政部门或专业学/协会发布的诊疗指导原则、指南等,选择相应评估指标。鉴于多数 AI 应用对不同级别医疗机构诊疗同质化程度有影响,可重点关注医联体或区域整体诊疗准确性和治疗成功率的变化影响。(6) 效率指标。主要反映 AI 医疗技术带来的临床工作流程优化和工作效率的改善,包括辅助诊断平均时间、平均手术时间、平均住院日、患者平均确诊时长等。例如 AI 医疗技术改善了患者自我管理现状,提高了用药依从性等。(7) 经济性指标。通常是测量并分析 AI 医疗技术的成本、效益、效果、效用等,综合判断 AI 医疗技术投入产出比,衡量其临床使用的经济价值及预算影响。政府角度开展的 AI 医疗技术经济性分析一般包含成本测算、成本费用差异分析、成本效果/效用/效益等增量分析;医院角度开展 AI 医疗技术经济性分析一般包括医院角度成本测算、运营分析。

3.3.5 数据安全 AI 需要使用大量多源数据,可能存在安全风险,如数据泄露、丢失等。因此,数据安全保护措施尤为重要,在 AI 临床应用评估时应包括数据安全性测评,如数据完整性、数据保密性、数据容灾能力等,具体可参考国家相关法律和规范,如《中华人民共和国数据安全法》《信息安全技术 健康医疗数据安全指南》《人工智能安全标准化白皮书(2023 版)》^[16-18]。此外,开发和应用 AI 医疗技术应确保训练和测试数据的准确性、均衡性、相关性、充分性、可靠性、规范程度符合要求;相关数据在最初选择、标注等环节不存在缺乏公平、包含偏见与歧视、范围与算法目标不匹配的问题;不存在被攻击或污染的情况。

4 试点研究

4.1 整体情况

在指南开发的同时,基于11家合作医疗机构,分批开展多场景的指南实证研究工作,包括目前已全部完成的冠心病AI辅助诊断系统(AI辅助冠状动脉CT血管成像(coronary computed tomography angiography, CCTA))卫生技术评估,以及正在实施中的人工智能肺结节CT影像辅助检测技术临床应用评估、智慧审方系统临床应用评估研究工作。

4.2 实证研究

以冠心病AI辅助诊断系统的卫生技术评估为例,根据指南提供的指导原则,包括以下流程。

4.2.1 明确技术类型 AI辅助CCTA技术属于本指南定义的一类技术,潜在风险属于三级,因此要进行临床应用评估以支持临床准入。

4.2.2 确定技术特性和临床应用场景 该技术算法和标准数据集已通过国家药品监督管理局医疗器械技术审评中心评审,并已在部分医院投入临床试用;预期用途主要是在CCTA图像后处理和生成报告阶段可以自动分割冠状动脉和评估冠状动脉狭窄程度;应用时的配套要求包括放射科医生、开具冠脉CTA检查的医生,配置塔式专用服务器、Ubuntu操作系统、23寸显示器以及病历AI终端智能调阅软件等,因此目前尚无法应用于社区卫生服务中心。冠脉CTA检查的患者和冠心病体检筛查人员是该项技术的主要适用人群。

4.2.3 确定待评技术的评估主题 该技术在临床试用阶段获得多数医生认可,但由于定价问题,无法正式进入临床。根据指南相关原则,选择有效性、安全性和经济性3个维度,评估使用与不使用该技术对医院运营产生的影响。

4.2.4 选择评估方法,具化评估指标 采用回顾性临床多中心真实世界研究设计,以42家医疗机构近6个月内接受数字减影血管造影术(digital subtraction angiography, DSA)介入检查或治疗的稳定性胸痛患者为研究对象,以传统的CCTA检查为

对照,对比分析AI辅助CCTA技术获取检查结果的时间效率及诊断准确性,以及因此而产生的成本和经济获益,最后建立决策树模型,模拟利用AI辅助CCTA技术对患者1年内的成本和效果产生的影响。

4.2.5 结果分析 经济学分析结果表明,AI辅助CCTA技术组人均总费用低于CCTA组,却获得了更大的临床效果,显示出该类AI技术的临床应用价值。同时验证了指南推荐的评估流程、方法和指标框架有较好可用性,指标框架操作性较好,核心数据可得且能得出相对客观的结果,反映出本指南具有实际应用参考价值。

5 结论与建议

AI临床应用评估是一项开拓性工作,旨在为此类创新技术的科学治理提供技术解决方案。本研究开发了首个针对AI临床应用评估的技术指南,形成了适用于我国AI临床应用评估的技术指导原则、方法、指标和基本流程,为指导开展规范评估、助力循证管理决策提供支持性工具;并通过多场景试点研究,证实了该指南的可操作性。结合实践发现未来仍待在以下领域深化研究。

一是进一步加强AI的临床应用场景定义和技术分类研究。医疗卫生是知识密集、技术密集行业,AI对传统医疗卫生诊治模式和思维形成冲击挑战,医疗卫生监管未能及时形成AI临床应用场景和功能的统一认知,造成技术治理理念滞后于新技术发展的问题,仍要组织AI、信息数据、临床等领域专家力量,针对AI临床应用场景进行深入定义和分析,形成既能体现技术本身分类特点、又能反映临床诊治规律的AI医疗技术分类标准,为科学治理政策制定奠定理论基础。二是持续强化评估指标体系研究构建工作。与传统诊疗技术相比,AI在跨机构和跨区域层面应用潜力更大,未来针对医联体和区域应用场景的评估指标仍要深入研究,特别是对于传统诊疗路径的优化、基层医疗机构赋能程度及各级各类医疗机构所提供服务的同质化影响等分析,仍要扩大现有指南应用实践,探索体系应用

场景下评估指标的选择,为深入挖掘有关创新技术的临床应用价值提供科学标准和技术工具。三是围绕专病领域建立标准指标集及数据集。真实世界证据将在 AI 医疗技术临床应用评估中发挥重要作用。专家咨询和试点研究均发现,不同医疗机构间指标变量定义、数据结构差异大,未来应逐步同质化核心基础变量和数据,探索形成不同病种标准评估指标集和数据集,规范 AI 医疗技术临床应用评估工作。四是加快研究建立评估质量控制体系,确保评估过程科学、评估结果可信。本次指南开发和试点工作初步探索了内外部质量控制标准,保障了研究结果的可靠性。未来仍要深入研究全国统一质量控制标准,针对各类评估方法的合理运用和结果的科学审核开发工具,为规范推动 AI 医疗技术临床应用评估工作、确保有关结果政策转化提供管理抓手。

参考文献

- 1 国务院. 新一代人工智能发展规划 [EB/OL]. [2023 - 09 - 05]. https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm.
- 2 国务院. “十四五”数字经济发展规划 [EB/OL]. [2023 - 09 - 05]. https://www.gov.cn/gongbao/content/2022/content_5671108.htm.
- 3 中共中央, 国务院. 数字中国建设整体布局规划 [EB/OL]. [2023 - 09 - 05]. https://www.gov.cn/xinwen/2023-02/27/content_5743484.htm.
- 4 邱陈辉, 黄崇飞, 夏顺仁, 等. 人工智能在医学影像辅助诊断中的应用综述 [J]. 航天医学与医学工程, 2021, 34 (5): 407 - 414.
- 5 上海交通大学人工智能研究院, 上海市卫生和健康发展研究中心, 上海交通大学医学院, 等. 中国人工智能医疗白皮书 [R]. 上海: 上海交通大学, 2019.
- 6 O' ROURKE B, OORTWIJN W, SCHULLER T. The new definition of health technology assessment: a milestone in international collaboration [J]. International journal of technology

- assessment in health care, 2020, 36 (3): 187 - 190.
- 7 CHALKIDOU K, MARTEN R, CUTLER D, et al. Health technology assessment in universal health coverage [J]. Lancet, 2013, 382 (9910): e48 - e49.
 - 8 陈洁, 于德志, 耿庆山. 卫生技术评估 [M]. 北京: 人民卫生出版社, 2013.
 - 9 国务院. 医疗器械监督管理条例 [EB/OL]. [2023 - 09 - 21]. https://www.gov.cn/premier/2021-03/20/content_5594138.htm.
 - 10 原国家卫生和计划生育委员会. 医疗技术临床应用管理办法 [EB/OL]. [2023 - 09 - 21]. https://www.gov.cn/gongbao/content/2018/content_5346680.htm.
 - 11 原国家卫生和计划生育委员会. 人工智能辅助治疗技术管理规范 [EB/OL]. [2023 - 09 - 21]. <https://www.chima.org.cn/Html/News/Articles/11000186.html>.
 - 12 国家药品监督管理局. 人工智能医疗器械质量要求和评价 [EB/OL]. [2023 - 09 - 21]. <https://www.nmpa.gov.cn/xxgk/ggtg/qtggtg/20220706104425173.html>.
 - 13 国家药品监督管理局. 人工智能医疗器械注册审查指导原则 [EB/OL]. [2023 - 09 - 21]. <https://www.cmde.org.cn/flfg/zdysz/zdyszwbk/20220309091014461.html>.
 - 14 原国家卫生和计划生育委员会. 人工智能辅助诊断技术管理规范 [EB/OL]. [2023 - 09 - 21]. <https://chima.org.cn/Html/News/Articles/11000184.html>.
 - 15 WHO. Ethics and governance of artificial intelligence for health: WHO guidance [EB/OL]. [2023 - 09 - 21]. <https://www.who.int/publications/i/item/9789240037403>.
 - 16 全国信息安全标准化技术委员会. 信息安全技术健康医疗数据安全指南: GB/T 39725 - 2020 [S]. 北京: 中国标准出版社, 2020.
 - 17 中华人民共和国数据安全法 [EB/OL]. [2023 - 09 - 05]. https://www.gov.cn/xinwen/2021-06/11/content_5616919.htm.
 - 18 全国信息安全标准化技术委员会. 人工智能安全标准化白皮书 (2023 版) [EB/OL]. [2023 - 09 - 05]. <https://www.tc260.org.cn/upload/2023-05-31/1685501487351066337.pdf>.