

基于卫生技术评估构建医学人工智能技术临床应用评估路径的探讨

赵锐^{1,2*} 石秀园³ 钟雪然⁴ 刘春平^{1,2} 任平^{1,2} 田雪晴^{1,2} 游茂^{1,2}

1. 国家卫生健康委卫生发展研究中心 北京 100044
2. 国家药物和卫生技术综合评估中心 北京 100044
3. 深圳市卫生健康发展研究和数据管理中心 广东深圳 518000
4. 广西医科大学 广西南宁 530021

【摘要】医学人工智能可以应用于疾病的预防、诊断和治疗等方面,以改善医疗服务的提供。卫生技术评估在国际上已广泛应用于新医药技术准入、医疗保险支付范围确定等方面。本文基于卫生技术评估,探讨构建我国医学人工智能临床应用评估路径,并提出了进一步发展的政策建议,为发展我国医学人工智能技术临床应用和卫生技术评估事业提供决策依据。

【关键词】卫生技术评估;医学人工智能;临床应用;评估路径

中图分类号:R197 文献标识码:A doi:10.3969/j.issn.1674-2982.2022.02.012

Enoploring of construction of clinical application assessment path on medical artificial intelligence technology based on health technology assessment

ZHAO Rui^{1,2}, SHI Xiu-yuan³, ZHONG Xue-ran⁴, LIU Chun-ping^{1,2}, REN Ping^{1,2}, TIAN Xue-qing^{1,2}, YOU Mao^{1,2}

1. China National Health Development Research Center, National Health Commission, Beijing 100044, China
2. National Center for Medicine and Health Technology Assessment, Beijing 100044, China
3. Shenzhen Health Development Research and Data Management Center, Shenzhen Guangdong 518000, China
4. Guangxi Medical University, Nanning Guangxi 530021, China

【Abstract】 Medical artificial intelligence can be applied to the prevention, diagnosis and treatment of diseases to improve the delivery of healthcare services. Health technology assessment has been widely used internationally in the access of new medical technology, the determination of medical insurance payment scope and so on. Based on health technology evaluation, this study discusses the construction of the evaluation path of clinical application of medical artificial intelligence in China and puts forward policy suggestions for further development, so as to provide a decision-making basis for the development of clinical application of medical artificial intelligence and health technology evaluation in China.

【Key words】 Health technology assessment; Medical artificial intelligence; Clinical application; Assessment path

人工智能(Artificial Intelligence, AI)最早于20世纪50年代在美国达特茅斯会议上被定义为“关于未来某种智能机器的设想”。世界卫生组织指出,人工智能系统是一种基于机器的系统,以不同程度的自主性学习,对于给定的一组人类定义的目标,进行推理、学习、适应、自我纠正、理解或互动,也可以影响真实或虚拟的决策。人工智能在医疗领域应用可

以改善医疗服务的可及性、质量和效率^[1],如可以帮助诊断、筛查或治疗,服务流程优化与资源分配,对患者进行依从性管理。但人工智能因其技术本身的特殊性,会对临床、人类认知、组织流程、经济、法律、伦理等多方面带来重大影响。随着科学和医学的进步,人工智能技术越来越多地被应用于临床领域,科学监管日益得到全球重视。我国目前按照《医疗器

* 基金项目:科技创新2030“新一代人工智能”重大项目(2020AAA0105000, 2020AAA0105001)

作者简介:赵锐(1986年—),女,博士,副研究员,主要研究方向为医药卫生政策评价、卫生技术评估。E-mail:zhaoruibsb@126.com

通讯作者:游茂。E-mail:ym@nhei.cn

械监督管理条例》对涉及到决策支持、辅助诊断的医用软件按照Ⅲ类医疗器械进行上市准入审批,即认为其具有较高风险,需要采取特别措施严格管控以保证其安全有效性。截至 2020 年底,已有 8 个国产和 1 个进口人工智能医疗器械通过审批并在国内上市应用,涉及类型包括眼科图像诊断、心电数据监测与诊断、生化数据监测和影像分析辅助诊断等,其中不乏附条件上市的产品。在资源有限投入的情况下,如何选择出符合临床需求的产品,并规范使用与管理,是政策决策者面临的重大难题。相比于传统的药品、器械耗材等卫生技术,人工智能作为一种颠覆性创新技术,其在原理、临床证据产生、监管方式等方面存在差异,需要对其评估路径进行设计与探讨。本研究在科技部重大专项的支持下,试图使用卫生技术评估(Health technology assessment, HTA)这一决策工具,建立对医学人工智能上市后临床应用评估的实施框架和路径。以期为我国更加科学、合理的开展医学人工智能上市后临床评价提供参考建议。

1 医学人工智能在临床应用中的潜在影响

1.1 人工智能在临床应用中的安全有效性问题

人工智能技术与传统医疗器械相比最大的特殊性在于它的研发和作用机制的不确定性和快速的技术迭代。这一特性需要市场准入监管者对上市后 AI 应用的安全有效性进行持续评估,通过获得更加丰富的临床高质量证据来验证其安全有效性。AI 临床应用最初的目的是辅助临床医生决策,但由于 AI 的应用多在实验室环境下,基于现实中已有且固化的数据,运用高质量的数据进行模型构建和决策训练,在真实诊疗场景下应用是否能与医疗专业人员水平保持一致受到质疑。^[2]如美国的一项研究表明,在前两家医院数据上训练的 AI 应用程序应用在第 3 家医院时并未得到预期效果。^[3]一些国家的临床医生认为 IBM Watson for Oncology 过于反映了美国治疗癌症的特殊性而放弃使用。^[4]同时,疾病和治疗方案是以非线性方式发展的,AI 技术的有效性呈现很大程度上取决于医生的技能和“学习曲线”。对于临床医生,其治疗方式和效果是随着经验愈加丰富而逐渐完善。人工智能的机器学习需要不断迭代其算法达到治疗有效,在监管中往往只需厂商证明这些微改变不影响原产品的安全性即可获批。而在监管过程中往往缺乏单个新产品高质量的临床试验数据,也缺乏同类产品之间可比较的数据,其效果的可评估性也是监管方的一个难题。

1.2 医学人工智能市场定价和临床应用的经济价值有待评估

人工智能开发和使用都需要较大的投入和支出,现实中的临床诊断和服务工作,需要整合各种类型的大量数据,需要专业的资源和费用来保证其进行持续的性能测试、数据质量安全、设备升级和专业人员培训。英国的一项研究评估了 AI 软件与标准治疗相比,其帮助检测脑卒中颅内大血管闭塞的成本效益。结果表明,AI 软件的使用可帮助减少 50% 的漏诊,在预期生命周期内可节省成本并可增加 QALY。^[5]其评价的成本效益是长期的,而软件成本是短期的。由于医疗人工智能的开发和引进对开发者和使用者来说费用较高,医疗机构对引进使用的动力不足。^[6]如对于已上市的医学影像人工智能系统,医院需要评价是否有必要花费较大成本引入此系统以减轻人工阅片的繁重工作。同时,我国大部分医疗人工智能的市场布局是为了提高基层医疗卫生机构诊断能力,促进上下级医疗机构之间的同质化。在基层的应用必然涉及医保购买以及群众支付意愿,需要进行预算影响分析来判断。

1.3 人工智能临床应用可能带来的伦理问题和对医患关系的挑战

人工智能临床应用对患者的数据隐私保护以及对传统医患关系的改变等,都对传统的临床诊断和医疗服务的质量监管带来新的伦理挑战。同时,AI 的应用可能会引发医生利用其机器学习的能力,诱导其学习更具有利益回报的诊断方式。如对较高的敏感度的临床阈值可能导致过度诊断或过度处方,而较低的敏感度会导致部分人群未被诊断和治疗。^[7]Wendy A 等人的一项研究对两个基于人工智能的临床决策支持系统(Painchek[®]和 IDx-DR)进行了特定环境和案例下的医疗伦理问题评估。^[8]Painchek[®]是辅助性的决策工具,为了提高如严重痴呆症患者的疼痛管理质量;IDx-DR 是自主的人工智能系统,用于自动检测糖尿病患者眼部疾病的早期迹象。该研究重点讨论了 AI 在诊疗过程中可能存在的欺骗行为、公众和患者参与(Public and patient involvement, PPI)以及医疗保健关系。评估结果显示,AI 存在潜在的欺骗性承诺风险,在医学 AI 的开发和部署中缺乏病人和公众的参与,并缺乏 AI 对医患关系影响的关注。关于训练集、算法的构建方式、敏感性和特异性、数据存储等方面的细节对于进行伦理评价至

关重要。每项医学人工智能应用都需要在其整个生命周期中进行审查。

2 基于卫生技术评估的医学人工智能临床应用评估路径探讨

2.1 卫生技术评估的制度化路径概述

有限的卫生资源与快速增长的健康需求之间的矛盾,是强化卫生技术评估应用于卫生决策工作的核心驱动。卫生技术评估是一种对卫生技术性质及影响的系统评估。^[9]主要目的是通过了解卫生技术对社会、经济、组织和伦理等方面产生的影响,为卫生政策制定提供高质量证据,确定一项新的卫生技术是否有更好的安全有效性、是否值得买,以及是否能够与目前服务体系相适应等(图1)。HTA 作为可支持多目标的战略政策工具可运用于卫生技术全生命周期的监管中,不论是卫生管理部门制定相关管理制度或医疗机构采购和管理,还是临床医生选择最佳诊断和治疗技术,不同类型 HTA 均能为其决策提供科学依据。HTA 在全球范围内广泛应用于支持药品、器械等医疗技术产品的政策制定、定价和报销决策,用于提高资源分配效率、提高服务质量和循证决策机制的建立,以应对由新卫生技术涌现、医疗需求增加等导致的医疗费用快速上涨和医保预算约束压力增大等引发的决策挑战。^[10]很多国家已经建立起较为完善的 HTA 制度体系,并提出了用于政策决策的制度化评估路径,包括四个主要阶段,即优先级设定、议题遴选、实施评估和审议推荐^[11-13],主要用于从政策议题的选定到转化为评估问题再用于提出政策建议的过程(图2)。使用 HTA 制度化评估路径将帮助相关 HTA 实施者更清晰地理解 HTA 制度化的政策进程并适时开展工作。

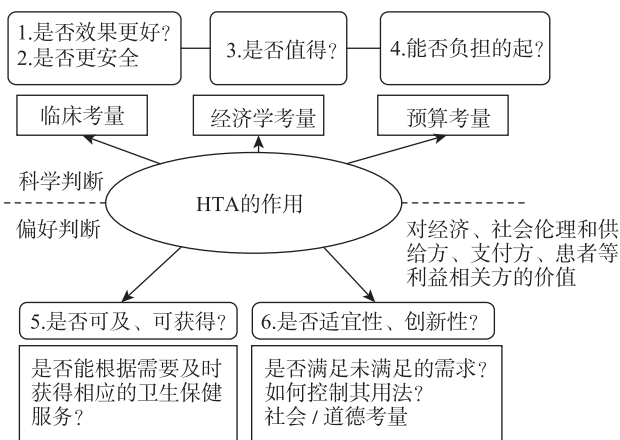


图1 卫生技术评估的作用

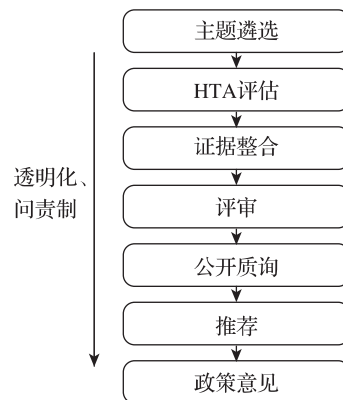


图2 卫生技术评估制度化路径框架

2.2 基于卫生技术评估的医学人工智能临床应用评估路径构建

《国务院办公厅关于改革完善医疗卫生行业综合监管制度的指导意见》(国办发〔2018〕63号)提出,强化国家卫生技术评估支持力量,发挥卫生技术评估在医疗技术、药品、医疗器械等临床准入、规范应用、停用、淘汰等方面的决策支持作用。随着 AI 技术在临床应用快速投入与更新,监管方、使用方和患者等各利益主体都面临着全新的挑战。与医学人工智能技术准入如火如荼地开展相比,其在选择和配置方面的科学管理还比较滞后。如相较于传统卫生评估, AI 最大的难点在于在不确定性和额外获益之间寻找平衡,同时能够兼顾多利益相关主体之间不同的愿景和目标,需要对医学人工智能技术的科学配置、规范管理和合理使用制定相应的管理和评估制度以及标准,确保上市后的安全、有效、经济可及。本文基于 HTA 制度化路径框架及相关领域专家建议,提出医学人工智能技术上市后在临床应用的评估路径框架(图3)。该路径以医学人工智能技术为评估对象,依次包括主题识别、主题遴选、技术评估、专家评审、质询转化五个基本步骤。

2.2.1 主题识别

主题识别是指对国内已经上市并在临床场景应用的人工智能技术产品进行收集、识别、分类和记录,并对其存在的政策问题形成主题清单的过程。这一过程需要依据优先级的排序结果将有限的卫生资源科学分配于有限的卫生技术评估资源。第一步,定期面向医疗卫生机构、相关行业协会等开展调研,或开展相关领域的专家咨询,收集并记录当前在临床场景应用的人工智能技术的具体信息。第二步,在官方网站上发布面向研发企业及机构的技术评估申报通知,有评估需求的企业及机构可通过网站申报技术产品的技术特性、临床应用场

景、应用表现等相关信息。第三步,将收集的以及企业申请的技术汇总,形成主题清单。每次完成主

题识别后,均把形成的主题清单内的主题纳入人工智能技术库。

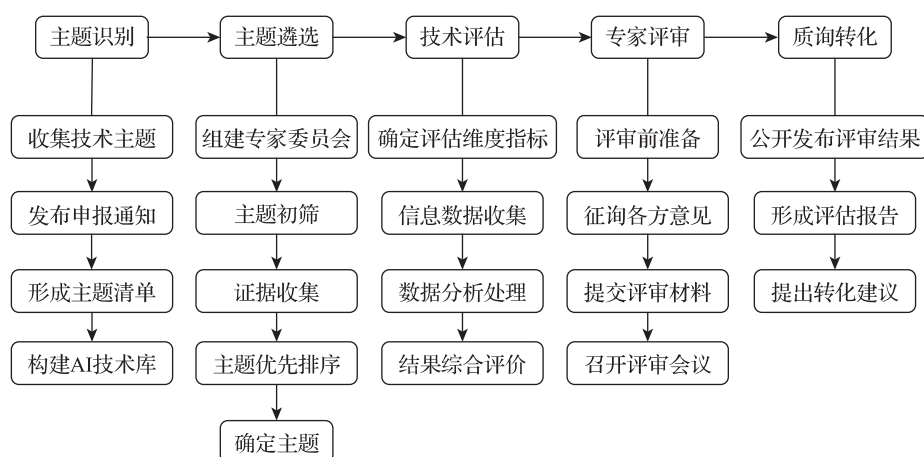


图3 医学人工智能技术临床应用评估路径框架

2.2.2 主题遴选

主题遴选是指在完成早期识别得到的主题清单基础上,以主题遴选专家委员会为依托,对清单进行证据收集、反复筛选,最终确定需要评估的技术和政策问题的过程。该工作可由卫生部门牵头定期开展,由下设主题遴选专家委员会和工作组实施具体工作。专家委员会成员应包括卫生决策者、人工智能相关领域专家、医务人员和患者等,主要负责主题筛选、对主题评分排序以及最终确定评估主题等工作。专家委员会依据疾病谱、疾病负担、现有 HTA 能力等因素,结合重点领域、具体决策需要、社会及公众关注度等,对主题清单进行初次筛选,形成初筛清单。遴选工作组围绕技术重要性、相关性、可评估性等维度,对各技术主题进行快速证据收集,形成技术主题信息摘要。接下来专家委员会基于信息摘要,对各技术主题进行评分、排序,形成主题积分排名表及综合评分主题排名表。最后,专家委员会根据主题信息摘要以及主题积分排名表,形成主题排序共识,最终确定本次评估的技术主题。

2.2.3 技术评估

技术评估指遴选出评估主题后,围绕具体维度及指标,对主题中每个技术产品进行数据收集、数据分析、综合评价的过程。技术评估步骤主要由评估工作组负责,形成关于该评估主题内不同技术产品的综合评价报告,是整个评估路径的核心步骤。首先,需结合具体评估需求及客观条件,确定评估需考虑的重点维度。在卫生技术评估中一般需考虑安全性、有效性、经济性、可及性、社会适应性等维度,根据确定的重点维度,甄选出每一个维度内评估所需的具体指标。其次,

评估工作组根据甄选出的评估具体指标,进行信息数据收集工作。信息收集方法通常包括文献研究、资料数据收集和填报、实地调研、座谈会、问卷调查等。在信息数据收集完成后,评估工作者需整理及进一步处理收集来的信息数据,对其进行科学的分析、分类和汇总,使其系统化、条理化、科学化,以反映所评估对象的特征和评估指标的要求。在数据分析得出结果后,不能将结果简单合并得出结论,应采取多维度决策分析法,全面考虑各个维度,通过运用定性及定量打分两种方式形成综合评价结果报告。

2.2.4 专家评审

专家评审是指将综合评价结果报告及其他相关证据进行整合评估,从临床、经济、社会和创新等维度判断一类临床应用的人工智能技术的综合价值,并作出最终决策意见的过程。评审工作应主要由人工智能技术评审专家组负责,工作组负责提供保障及辅助工作。首先由国家或省级评估中心通知评审专家组启动评审工作程序。第二步,评估工作组通过召开利益相关方的意见征询会,包括行业协会、企业代表、医护人员及患者代表等,聆听评估结果并提供反馈意见,评估工作组负责记录利益相关方的意见建议并整理形成意见汇总表。工作组将整理好的相关评审材料(结果综合评价报告、利益相关方意见汇总表等)发送给各评审组专家,帮助评审组专家深入了解评估项目。最后,评审专家组召开闭门会议,审议并讨论相关材料,形成一致的专家评审意见及政策建议。对于争议大、证据不足的评估结果,评审专家组组长有权利在征求组内意见后向评估实施方提出继续完善评估的建议。

2.2.5 质询转化

质询指公开发布评估报告及评审意见,收集有关意见建议。转化指基于质询后的最终评估报告,制定涉及监管审批、规划配置、临床应用等政策建议。在评审工作完成后,评估工作组将评估报告扩展版摘要与评审意见一并发布于评估实施方的官方网站上,全面征求有关各方意见,并判断是否需要进行评估报告的调整。评估报告如有重大调整,或将对评审意见产生影响,评估实施方将上报卫生决策部门,在评估报告再次完善后需启动二次专家评审程序,直至形成最终的技术评估报告。评估报告正式定稿后,在国家或省级卫生健康行政部门官方网站发布。最后,评估实施方应结合相关政策文件,研究制定促进结果转化应用的措施及建议。具体转化路径包含但不限于:临床指南的制定与调整、循证知识传播、基本服务包制定、定价与补偿制度完善、医疗行为监管强化等。

3 实施医学人工智能临床应用评价的相关建议

人工智能技术在健康领域的应用势不可挡,基于 HTA 构建的医学人工智能临床应用评估路径,为更好理解和运用 AI 技术的特征,开展更加科学灵活的监管提供了一个可供参考的思路。在具体的实践中,还需要国家从以下方面进行顶层设计。

一是借鉴相关领域经验,加快推进医学人工智能临床应用评估组织体系和机制建设。目前我国在人工智能应用、评估探索方面已走在世界前列,但医学人工智能技术临床应用评估体系还未建立。可借鉴药品、医疗器械等已有评估制度的经验,加强我国 HTA 应用于医学人工智能评估的组织管理体系建设,加强评价人才队伍培养,发挥国家级 HTA 中心和地方 HTA 中心的力量。配套完善相关法律法规,为评估体系应用提供指导与规范,更好地防控人工智能技术带来的潜在风险,保障评估工作的美好有序开展,为人工智能全生命周期监管提供有效助力。

二是建立多部门参与的评价研究和协作机制。人工智能技术的利益相关方要考虑到 AI 技术的特殊性,以及不同利益相关者的利益考量,对同类型的研究证据进行不同的权衡和取舍,这就需要强化政府主导和部门协作,构建政府、专业机构、学术团体、企业和公众等多方参与的研究和转化平台,提高评价的科学性和公平性。^[14]

三是加强评估结果的转化应用,推动评价结果服务于政策决策。重视医学人工智能技术上市前审

批与上市后评估的相互结合。评估结果可用于医院准入、医保购买、患者支付、创新技术替代等。

作者声明本文无实际或潜在的利益冲突。

参 考 文 献

- [1] He J, Baxter S L, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine[J]. *Nature Medicine*, 2019, 25(1): 30-36.
- [2] Panch T, Mattie H, Celi L. The “inconvenient truth” about AI in healthcare[J]. *npj Digital Medicine*, 2019, 2: 77
- [3] Zech J R, Badgeley M A, Liu M, et al. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study[J]. *PLoS Medicine*, 2018, 15(11): e1002683.
- [4] Ross C, Swetzlitz S I. IBM Pitched Its Watson Supercomputer as a Revolution in Cancer Care. It's Nowhere Close [N]. *STAT*, 2017-09-05.
- [5] van Leeuwen K G, Meijer F J A, Schalekamp S, et al. Cost-effectiveness of artificial intelligence aided vessel occlusion detection in acute stroke: an early health technology assessment[J]. *Insights into Imaging*, 2021, 12(1): 133.
- [6] 蒋璐伊,王贤吉,金春林. 人工智能在医疗领域的应用和准入[J]. *中国卫生政策研究*, 2018, 11(11): 78-82.
- [7] Alami H, Lehoux P, Auclair Y, et al. Artificial Intelligence and Health Technology Assessment: Anticipating a New Level of Complexity[J]. *Journal of Medical Internet Research*, 2020, 22(7): e17707.
- [8] Rogers W A, Draper H, Carter S M. Evaluation of artificial intelligence clinical applications: Detailed case analyses show value of healthcare ethics approach in identifying patient care issues[J]. *Bioethics*, 2021, 35(7): 623-633.
- [9] WHO. Health technology assessment [EB/OL]. (2014-09-02) [2021-12-01]. https://www.who.int/medical_devices/assessment/zh/
- [10] 吕兰婷,傅金澜. HTA 决策转化国际对比研究及经验借鉴: 基于 HTA 十五项基本原则[J]. *中国药物经济学*, 2019, 14(6): 113-118.
- [11] 茅艺伟,陈英耀,唐檬,等. 澳大利亚卫生技术评估的应用[J]. *中国卫生资源*, 2014, 17(6): 484-486.
- [12] 严俊涛,魏艳,刘世蒙,等. 国际框架对我国卫生技术评估制度化路径构建策略的启示[J]. *中国医疗保险*, 2022(1): 75-80.
- [13] NICE E. Guide to the processes of technology appraisal [EB/OL]. (2014-09-02) [2021-08-12]. <https://www.nice.org.uk/process/pmg19/chapter/acknowledgements>
- [14] 李海龙,徐佩佩,曾力楠,等. 卫生技术评估: 医疗器械科学监管的重要工具[J]. *中国循证医学杂志*, 2021, 21(1): 2-7.

[收稿日期:2021-11-03 修回日期:2022-01-10]

(编辑 赵晓娟)