• 医学信息研究 •

我国医疗人工智能风险研究现状及风险识别*

董 怡[1] 冉 晔[1] 余中光[2]

([1] 北京中医药大学人文学院 北京 102488 [2] 中日友好医院呼吸中心 北京 100029)

〔摘要〕 **目的/意义** 梳理我国医疗人工智能风险研究现状，针对相关风险提出应对策略。**方法/过程** 运用文献计量学方法识别研究热点。基于技术－组织－环境理论和文本分析法，从环境、组织、技术、个体 4 个层面总结并提出医疗人工智能风险识别理论框架。**结果/结论** 共识别出 24 项具体风险，并从制度体系建设、组织风险管控、个体义务范围和技术应用等方面提出建议。

〔关键词〕 医疗人工智能；技术－组织－环境理论；风险识别

〔中图分类号〕 R－058 〔文献标识码〕 A 〔DOI〕 10.3969/j.issn.1673－6036.2025.03.005

Study on the Current Status of Medical Artificial Intelligence Application Risk Research and Risk Identification in China

DONG Yi[1], *RAN Ye*[1], *YU Zhongguang*[2]

[1]*College of Humanities, Beijing University of Chinese Medicine, Beijing 102488, China*; [2]*Respiratory Centre, China－Japan Friendship Hospital, Beijing 100029, China*

〔**Abstract**〕 **Purpose/Significance** To sort out the current status of medical artificial intelligence (AI) risk research in China, and to propose coping strategies for related risks. **Method/Process** Research hotspots are identified by using bibliometric methods. Based on the technology－organization－environment theory and textual analysis method, the paper summarizes and proposes a theoretical framework for medical AI risk identification from 4 levels: environment, organization, technology, and individual. **Result/Conclusion** A total of 24 specific risks are identified and recommendations are made in terms of institutional system construction, organization risk control, scope of individual obligations, and technology application, etc.

〔**Keywords**〕 medical artificial intelligence; technology－organization－environment theory; risk identification

1 引言

医疗人工智能（artificial intelligence，AI）指使用 AI 技术处理、分析和理解医疗数据，以改进医疗服务和健康管理的学科及技术领域。随着 AI 技术的逐步发展，其已被广泛应用于医学影像分析、个性化医疗、辅助诊断、健康监测与管理、疾病预测与预防等多个领域[1]。然而，医疗 AI 在应用过程中面临诸多风险，精确识别是有效应对的前提。2024 年 3 月 13 日欧洲议会通过《人工智能法案》，

〔修回日期〕 2024－11－22

〔作者简介〕 董怡，硕士研究生；通信作者：余中光。

〔基金项目〕 国家自然科学基金青年基金项目（项目编号：72104255）；中国医学科学院医学与健康科技创新工程项目（项目编号：2021－I2M－1－046）；中日友好医院高水平项目（项目编号：2024－NHLHCRF－GL－12）。

该法案以预防风险为原则，引入多级风险监管框架[2]。我国医疗 AI 风险治理尚缺乏统一标准和规范。因此，本文从文献计量学视角回顾和分析过去 10 年间我国医疗 AI 风险研究现状，基于精读文献进行风险识别，为风险应对提供参考。

2 研究方法及研究现状

2.1 研究方法

运用文献计量学方法，在中国知网检索学术期刊，分别以“医疗 + 人工智能 + 风险”“医学 + 人工智能 + 风险”为主题组配检索，时间范围设为 2015—2024 年，检索时间为 2024 年 5 月 15 日，共检出相关文献 368 篇。剔除重复和非学术文献，最终纳入定性分析文献 230 篇。以 EndNote 格式导入 VOSviewer 和 UCINET 软件[3]，选择关键词聚类进行数据挖掘，绘制可视化知识图谱，呈现研究趋势及热点。运用文献分析法，精读 230 篇文献，最终筛选出医疗 AI 风险识别相关研究 43 篇，提炼年份、关键词、摘要、核心观点等，分析风险要素点，完成影响因素和具体风险点的归类统计。以技术 - 组织 - 环境（technology - organization - environment，TOE）理论为框架，纳入个体因素，从环境、组织、个体和技术维度开展风险识别。

2.2 现状分析

2.2.1 年度发文量分析 医疗 AI 风险相关文献年度发文量，见图 1。我国医疗 AI 风险研究起步较晚，2016 年起稳步增长，2018 年起快速增长。2018—2020 年，医疗 AI 风险研究引发学界广泛关注，发文量迅速上升，2020—2022 年稍有回落，2023 年达高峰。

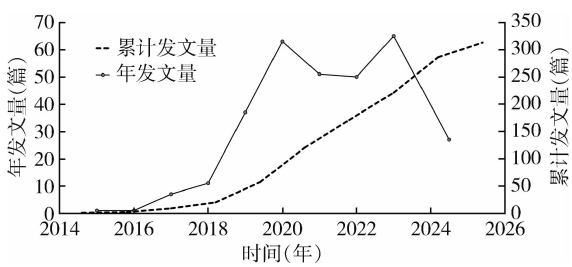

图 1 医疗 AI 风险相关文献发文趋势

2.2.2 聚类分析 从 230 篇相关文献中提取关键词共 789 个，高频关键词主要包括机器学习、大数据、智慧医疗、医患关系、个人信息保护等。词频大于等于 2 的关键词 128 个，运用 VOSviewer 可视化软件聚类，见图 2。共生成 128 个节点、368 条连线，得到 3 个聚类。聚类一：医疗 AI 伦理与法律风险，探讨伦理和法律在医疗 AI 应用中的角色，对其产生的伦理风险[4]及责任界定[5]问题进行深入研究，强调伦理规制[4]和法律保护的必要性。聚类二：技术与数据安全风险，聚焦技术风险研究、疾病预测诊断系统的研发及应用[6-8]等。聚类三：风险治理研究，主要聚焦于风险评估、风险预警、风险治理、风险预测等。有学者[9]提倡摒弃传统治理思维，在医疗 AI 中厚植人文关怀。

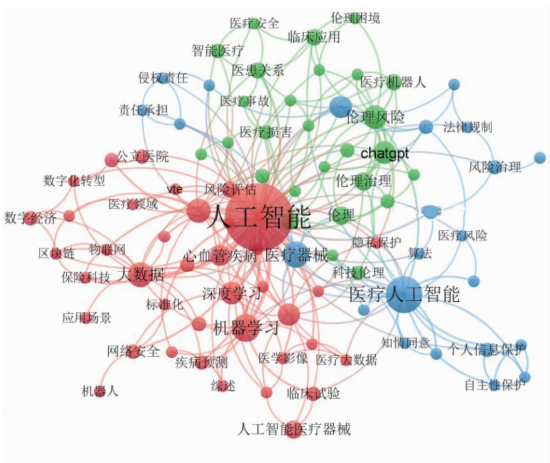

图 2 医疗 AI 风险相关文献关键词聚类

3 TOE 理论框架下医疗 AI 风险识别

3.1 TOE 理论框架

TOE 理论[10]将影响企业技术创新实施的因素归纳为技术、组织和环境 3 个维度。随着相关研究的增多，其内涵和外延不断丰富。本研究基于该理论，从 230 篇文献中筛选 43 篇精读，结合聚类图谱、文献梳理及实务案例，鉴于医疗 AI 所连接的医患双方个体特性，纳入个体维度，提炼归纳医疗 AI 领域具体风险 24 项，构建医疗 AI 风险识别理论框架，见表 1。

表 1 医疗 AI 风险识别理论框架

维度	影响因素	具体风险
环境	政策与法规	权责划分不明①、伦理规范缺失
	社会环境	资源分配不均、就业危机、医学人文淡漠
组织	医疗机构	医患纠纷、组织信任度、患者满意度、数据管理、数据准确性
个体	医务人员	自主性弱化、规范操作、机能退化、职业倦怠、医患关系
	患者	就医成本、知情同意、隐私保护、侵犯人权尊严、生理伤害②
技术	—	机器误差、技术衔接、技术依赖、缺乏共情

注:①参见北京市第一中级人民法院（2021）京 01 民终 8552 号民事判决书;②参见广东省佛山市南海区人民法院（2022）粤 0605 民初 14328 号民事判决书。

3.2 环境维度风险

环境维度聚焦于外部政策法规和社会环境，具体风险表现为医疗权责划分不明、伦理规范缺失、资源分配不均、就业危机、医学人文淡漠等。医疗 AI 具有“黑箱”属性，其责任主体、监管模式较难明确。许多国家或国际组织出台指导原则，如欧盟《人工智能法案》《医疗人工智能：应用、风险以及伦理与社会影响》，美国《人工智能风险管理框架》等[11]。我国虽出台《新一代人工智能伦理规范》和《关于加强科技伦理治理的意见》[12]，但政策体系不够细化[13]，且暂未出台专门聚焦于医疗卫生领域的伦理规范制度，绝大部分伦理准则、规范和制度仍碎片化分布于不同的政策和法律文件中，亟需建立统一、细化的医疗 AI 伦理规范制度[14]。医疗 AI 应用带来资源分配不均问题，数据和算法可能导致垄断和偏见，受益群体集中在小部分发达地区或富裕人群，加剧医疗资源分配不均现象。随着医疗 AI 在智能诊疗领域的应用场域不断拓展，AI 具备的强大学习和精准医疗能力可能给传统医师带来失业风险。同时，医疗 AI 技术对医疗专业人员技能要求较高，可能造成医疗行业人才结构不合理，加剧就业不平衡问题。医疗 AI 的应用冲击了传统医患交流模式，可能导致医学人文关怀的弱化趋势。

3.3 组织维度风险

组织维度关注医疗机构内部的组织结构和组织管理。医疗 AI 应用存在医患纠纷增多及信任度、满意度降低等风险。医疗 AI 决策与医务人员的判断可能不一致，医生负有再判断义务[15]，易造成医疗资源重复浪费。医疗 AI 算法和模型依赖大量医疗数据，需关注数据质量和安全，目前医疗机构和技术提供商对于医疗数据的所有权边界不明确，产生诸多问题。此外，数据质量不佳导致诊疗结果不准确和算法偏见[1]。技术不合理应用产生依赖性，医疗 AI 技术的高效便捷可能导致医生“警觉衰退”[16]。此外，一旦医疗 AI 系统受到攻击，严重影响数据安全，增加医患纠纷，影响医疗机构声誉，降低患者信任度和满意度。

3.4 个体维度风险

3.4.1 *医方层面* 医疗 AI 应用场域不断拓展，在实际应用中会产生医患关系恶化、医生主体地位被削弱[17]、机能退化及职业倦怠等风险。在医疗 AI 辅助医疗决策的过程中，医生容易过度依赖 AI，从而减弱自身医疗决策能力，长此以往，会导致医生个体层面机能退化。机器判断并非绝对准确，医生负有再判断义务[15]，加剧了医生诊疗负担，且机器判断极易干扰医生独立判断。新技术和新设备的引入会增加医务人员技术操作负担，加剧医务人员职业倦怠风险。医疗 AI 应用还可能导致医患信息不对称、沟通不足等问题，影响治疗效果或加剧医患关系紧张。

3.4.2 *患方层面* 医疗 AI 应用使就医流程及模式复杂化，增加就医成本。目前大部分医疗机构对医疗 AI 告知仍不规范，医生多不主动告知患者 AI 参与情况，易侵犯患者知情同意权。医疗 AI 前期准入阶段，可能存在未经批准注册的应用，威胁患者生命健康。在临床应用过程中，医疗 AI 辅助诊疗可能会出现误诊、漏诊或系统故障等情况，对患者造成生理、心理创伤，甚至危及生命。此外，技术冰冷性易导致医患缺乏共情，侵犯患者的尊严和人权[5]。医疗 AI 处理患者数据时存在泄漏风险，如未经授权访问或安全漏洞，将侵犯患者隐私权。

3.5 技术维度风险

技术维度强调技术兼容性、医疗数据质量与安全、算法决策公正透明。医疗 AI 发展得益于机器学习、算法决策、海量数据处理等[17]。不同医疗机构所使用的医疗 AI 系统可能受数据质量、算法设

计或模型过拟合影响，产生机器间误差，导致诊断结果不一致。医疗 AI 系统需与医疗设备、电子病历系统等医疗信息系统衔接，若技术衔接或信息传递不畅，可能导致医疗信息延误、丢失或错误。

4 讨论

4.1 研究意义

本文运用文献计量学及文本分析法，系统梳理 2015—2024 年我国医疗 AI 风险相关文献，概括研究现状和热点。TOE 理论为医疗 AI 风险识别提供全面系统的分析视角，助力理解和应对医疗 AI 技术创新带来的风险。本文在 TOE 理论[10]环境、组织和技术 3 要素的基础上，纳入医患双方个体要素，提出医疗 AI 风险识别理论框架，阐明 24 类具体风险，拓宽 TOE 理论结构和层次，为后续进一步提出风险评估指标提供理论支撑。此外，TOE 理论框架为本研究提供了开放动态的思维模式，适应医疗 AI 不断发展变化的内外部环境要求。后续研究可在此基础上细化研究风险评估指标，建立主观评价与客观资料、数据相结合的新型风险评估方法，为风险管理决策者确定合理的风险收益比提供证据支持，最大限度减少人为主观因素，实现 AI 风险识别预警，推动医疗 AI 安全应用与可持续发展。

4.2 启示和建议

4.2.1 完善医疗 AI 社会层面制度体系建设 在医疗 AI 开发和应用阶段，加强法律、伦理和技术规范的前期规制。一是针对个人信息和隐私立法不完善[15]及责任主体模糊、“多因一果”导致的责任界定不明确等问题，立法应植入数据保护理念，强化监管者和设计者责任，从增设罪名、罪状方面入手，明确事故归责原则并厘清责任权属[1]，注重各部门法之间的贯通。鉴于医疗 AI 的多变性和复杂性，应加大监管投入，探索建立灵活监管框架[18]。二是针对我国政策体系规定不细化[19]及伦理规范制度碎片化现状，亟须建立统一细致的医疗 AI 伦理规范制度。三是为指导医疗 AI 政策体系落地，制定实践指南必不可少。2021 年世界卫生组织发布《卫生领域 AI 的伦理与治理》[20]指南，我国可基于国情，借鉴参考其先进经验。指南制定需医务工作者、医疗 AI 专家、伦理学专家、法律专家多方参与，加强跨学科交流[21]。内容方面，指南应明晰多主体的责任和义务，完善科技伦理标准。完善的法律规范体系能够保障指导原则融入医疗 AI 全生命周期。

4.2.2 多举措提升组织层面风险管控能力 一是医疗机构可设立专门部门，负责制定医疗 AI 风险管理政策和流程[22]，监督和指导技术实施。医务人员应参与医疗 AI 规则设计与开发，结合实际临床流程，避免技术设计与实际临床工作流程不符。二是医疗机构在引入医疗 AI 前，应为设备与数据管理提供制度保障，建立全面的风险评估机制，包括制定科学的评估办法、设计可行的风险评估指标、探索建立风险评估动态更新库等[14]，充分考虑医疗 AI 的适用性及风险，并与产品服务提供商明确各自权责。三是在应用医疗 AI 过程中，医疗机构应构建规范的管理模式，包括告知制度、数据管理制度、考核制度和设备异常情况反馈制度。

4.2.3 明确医务人员的责权范围 医务人员作为医疗 AI 应用的关键一环，首先应积极顺应医疗 AI 发展趋势，不断强化自身参与度。其次，应主动学习医疗 AI 系统原理、操作规范、技术伦理等相关知识，通过参加培训或自学课程，提高医疗 AI 技术应用水平，实现医疗 AI 临床规范化运用。在技术发挥便捷性的同时，医务人员应保持敏锐，审慎对待医疗 AI，避免过度依赖。再次，应注重与患者保持良好沟通[22]，向患者通俗化解释医疗 AI 的诊疗参与情况和风险，充分保障患者知情权和自主选择权，促进医患开放对话，消除患者对医疗 AI 应用的顾虑和紧张情绪，促进医患信任。

4.2.4 构建医疗 AI 技术伦理合规性与良性发展协同机制 基于技术正当性理论框架构建算法透明化体系和数据治理体系。在技术研发阶段，应秉持公平正义的价值导向[23]，通过医疗机构异常数据回溯分析与临床反馈，实现技术迭代与伦理准则的动态耦合。在技术运行过程中，应恪守程序正当原则，采取技术手段规制算法。为了更好地监测技术的运行，可在医疗 AI 中嵌入监测系统，自动监测和报告错误。数据是医疗 AI 发展的基础，数据治理方面应构建三重协同框架。其一，完善全生命周

期数据安全防护体系，提升数据存储系统的稳定性和抵御黑客攻击的能力[24]。其二，推进医－研－企数据共享机制建设，通过标准化数据清洗流程与可解释性增强技术降低误差风险。其三，优化不同医疗信息系统之间的衔接和数据交互技术，更好地发挥医疗大数据优势，促进医疗 AI 研发。

5 结语

本研究仅基于中国知网检索期刊文献，未纳入其他数据库和其他类型文献，存在局限。后续将扩大数据库检索范围，加强国内外文献资料收集整理，以获得更加科学严谨的结果。研究方法方面，本研究基于 TOE 理论分析框架，初步提出医疗 AI 风险识别理论框架，但研究方法相对单一。后续应通过访谈、实地观察、扎根理论等深入识别风险，完善我国医疗 AI 风险识别理论框架及应对策略。

作者贡献：董怡负责文献调研、数据分析、论文撰写；冉晔负责论文修订；余中光负责研究设计。
利益声明：所有作者均声明不存在利益冲突。

参考文献

1 徐着雨，岳远雷．医疗人工智能算法风险防范的法治化思考［J］．医学与哲学，2023，44（11）：67－71.
2 陈建兵，王明．负责任的人工智能：技术伦理危机下 AIGC 的治理基点［J］．西安交通大学学报（社会科学版），2024，44（1）：111－120.
3 吕静，何平，王永芬，等．ChatGPT 在医学领域研究态势的文献计量学分析［J］．医学与哲学，2024，45（7）：30－35.
4 张荣，徐飞．智能医学算法决策的伦理困境及风险规制［J］．医学与哲学，2022，43（8）：10－15.
5 石佳友，徐靖仪．医疗人工智能应用的法律挑战及其治理［J］．西北大学学报（哲学社会科学版），2024，54（2）：91－103.
6 刘霞，石元伍，王文聪，等．基于智慧物联网的社区老年健康监测服务设计研究［J］．包装工程，2024，45（14）：128－136.
7 GOSAK L，MARTINOVIĆ K，LORBER M，et al. Artificial intelligence based prediction models for individuals at risk of multiple diabetic complications：a systematic review of the literature［J］. Journal of nursing management，2022，30（8）：3765－3776.
8 曹敦煜．人工智能在心脏疾病诊疗中的应用［J］．科技传播，2019，11（4）：141－142.
9 徐明，韦俨芸．数字时代医疗人工智能的算法逻辑、风险及其应对［J］．中南民族大学学报（人文社会科学版），2024，44（1）：146－154，187.
10 段淳林，崔钰婷．颗粒度、信息质量和临场感：计算广告品牌传播的新维度——基于 TOE 理论的研究视角［J］．武汉大学学报（哲学社会科学版），2022，75（1）：79－90.
11 张浩．人工智能治理的实践进展与展望［J］．人工智能，2022（1）：16－21.
12 张吴越，周庆，桑爱民．医院伦理审查委员会建设的新思考［J］．中国医学伦理学，2022，35（9）：986－989，1006.
13 薛澜，贾开，赵静．人工智能敏捷治理实践：分类监管思路与政策工具箱构建［J］．中国行政管理，2024，40（3）：99－110.
14 黄成华．智慧医学的意义解读、风险研判及其防范路径［J］．中国医学伦理学，2023，36（12）：1313－1322.
15 吴何奇．大数据时代个人隐私保护的刑法路径——从医疗人工智能的隐私风险谈起［J］．科学与社会，2020，10（2）：89－110.
16 谭璐，刘小红．“人－机”医疗模式下的伦理学问题及应对策略［J］．中国医学伦理学，2019，32（9）：1127－1131.
17 戎春宇，洪冬旎，王宝悦，等．人工智能在公共卫生领域研发应用的伦理思考［J］．上海预防医学，2024，36（5）：504－510.
18 王颖．我国数字医疗监管原则在独立软件医疗器械的应用初探［J］．药学研究，2024，43（4）：359－365，395.
19 郑志峰．诊疗人工智能的医疗损害责任［J］．中国法学，2023（1）：203－221.
20 MANICKAM P，MARIAPPAN S A，MURUGESAN S M，et al. Artificial intelligence（AI）and internet of medical things（IoMT）assisted biomedical systems for intelligent healthcare［J］. Biosensors，2022，12（8）：562.
21 HERINGTON J，MCCRADDEN M D，CREEL K，et al. Ethical considerations for artificial intelligence in medical imaging：data collection，development，and evaluation［J］. Journal of nuclear medicine，2023，64（12）：1848－1854.
22 SMITH H，FOTHERINGHAM K. Artificial intelligence in clinical decision－making：rethinking liability［J］. Medical law international，2020，20（2）：131－154.
23 徐娟．智慧医疗运行风险防控的法治化策略［J］．甘肃政法大学学报，2024（1）：88－99.
24 王玥，宋雅鑫，王艺霏，等．卫生领域人工智能的伦理与治理：多模态大模型指南［J］．中国医学伦理学，2024，37（9）：1001－1022.