

DOI: 10.16781/j.CN31-2187/R.20250127

• 专题报道 •

人工智能在医疗决策中的法律责任与伦理挑战：基于诊疗场景的分析

王圆明¹, 葛俊骁², 徐青松^{2*}

1. 解放军总医院京北医疗区红山门口门诊部, 北京 100091

2. 海军军医大学(第二军医大学)基础医学院, 上海 200433

[摘要] 当人工智能融入医疗决策时, 技术自主权的模糊界定与责任分配的多主体博弈催生出法律权责争议与伦理治理难题。本文以临床诊断、治疗策略生成及预后评估为焦点, 揭示算法“黑箱”引发的数据偏倚与决策可溯性缺陷, 并从责任主体界定、技术缺陷标准化评估及跨学科协同三向度提出解决方案。在法律维度上, 建议基于欧盟《人工智能法案》的“高风险系统”分类与产品责任相关法规的无过错追责原则, 构建开发者-操作方-医疗机构三级责任框架, 即辅助型系统由操作医师承担终局责任, 决策型系统则强制开发者提供可验证算法并承担连带责任。在伦理层面上, 建议通过动态审查机制(如强制公开决策树逻辑与数据均衡性指标)平衡技术效能与患者知情权。

[关键词] 人工智能; 医疗决策; 法律责任; 伦理风险; 算法透明度

[引用本文] 王圆明, 葛俊骁, 徐青松. 人工智能在医疗决策中的法律责任与伦理挑战: 基于诊疗场景的分析[J]. 海军军医大学学报, 2025, 46(8): 977-981. DOI: 10.16781/j.CN31-2187/R.20250127.

Legal liability and ethical challenges of artificial intelligence in medical decision making: analysis based on diagnosis and treatment scenarios

WANG Yuanming¹, GE Junxiao², XU Qingsong^{2*}

1. Red Hill Outpatient Clinic, Northern Medical Branch of Chinese PLA General Hospital, Beijing 100091, China

2. College of Basic Medical Sciences, Naval Medical University (Second Military Medical University), Shanghai 200433, China

[Abstract] When artificial intelligence is integrated into medical decision-making, the ambiguous definition of technological autonomy and the multi-party game of responsibility allocation lead to legal disputes over rights and responsibilities as well as ethical challenges. Focusing on clinical diagnosis, treatment strategy generation, and prognosis assessment, this paper reveals the data bias and decision traceability deficiencies caused by the “black box” of algorithms, and then proposes solutions from 3 dimensions: defining responsible parties, standardizing technical defect assessments, and promoting interdisciplinary collaboration. From a legal perspective, based on the “high-risk system” classification under the European Union *Artificial Intelligence Act* and no-fault liability principle under the product liability relative laws, a 3-tier responsibility framework of developer-operator-medical institution is constructed: for auxiliary systems, operating physicians bear ultimate responsibility; for decision-making systems, developers are required to provide verifiable algorithms and take joint liability. Ethically, dynamic review mechanisms (such as mandatory disclosure of decision tree logic and data balance metrics) are suggested to balance technological efficiency with patients’ right to informed consent.

[Key words] artificial intelligence; medical decision; legal liability; ethic risk; algorithm transparency

[Citation] WANG Y, GE J, XU Q. Legal liability and ethical challenges of artificial intelligence in medical decision making: analysis based on diagnosis and treatment scenarios[J]. Acad J Naval Med Univ, 2025, 46(8): 977-981. DOI: 10.16781/j.CN31-2187/R.20250127.

人工智能(artificial intelligence, AI)技术对医疗决策的深度渗透正重塑传统诊疗范式, 其作为核心驱动力已贯穿诊断、治疗策略生成及预后评估等重要环节。值得关注的是, AI在提升诊疗效率与精准度的同时催生多维度权责冲突与伦理困境,

集中表现为技术自主性边界模糊、责任分配机制交织及算法“黑箱”引发的决策可溯性缺陷^[1]。在临床医疗决策场景中, 智能系统不仅重构了医患互动模式, 也给责任认定、数据隐私保护与伦理审查机制等方面带来了新的挑战。在技术革新与风险管

[收稿日期] 2025-03-05 [接受日期] 2025-06-04

[作者简介] 王圆明, 硕士生. E-mail: 178245792@163.com

*通信作者(Corresponding author). E-mail: nils_2003@163.com

控间寻求动态平衡并构建标准化的AI医疗决策框架,对于促进医疗AI健康发展和实践应用具有重要意义。

1 AI在医疗领域的应用现状

1.1 医疗保健的智能化决策支持 AI技术在医疗保健中的应用正由基础操作层面面向决策支持领域延伸。在生命体征动态监测中,AI系统通过对患者心电波形与血压变化曲线时序数据进行分析,能够精准识别异常生理波动并触发预警机制,为医疗保健团队提供科学决策依据^[2]。医学影像智能解析技术可辅助医护人员快速定位关键病理征象,同时结合自然语言处理技术提取电子病历中的核心症状特征,显著提升医疗干预的时效性^[3]。基于多模态健康数据的个性化镇痛决策模块通过药物代谢基因组学特征优化给药方案,有效控制药物不良反应发生风险^[4]。AI技术已深度融入医疗保健决策全流程,但其对医护人员专业判断权的潜在影响仍需通过系统性评估加以规范。

1.2 医学诊断的精准化决策革新 AI技术正驱动医学诊断由经验主导转向数据驱动范式,其核心效能在于增强决策精准度与效率。在影像诊断领域,基于深度学习的辅助系统可捕捉早期微小病灶并为医生输出决策数据^[5]。在病理分析中,AI通过高维数据处理技术能够从海量切片中锁定并识别肿瘤演进关键标志物,帮助病理医师优化诊断流程^[6]。在临床数据整合层面,通过集成病史、体征及实验室检测结果,AI系统可生成多维诊断矩阵,使医院误诊率明显降低^[7]。值得注意的是,基于5G技术的远程诊断平台,AI技术可使专家决策支持覆盖率显著提升,但过度依赖算法导致的临床决策能力弱化风险仍需警惕^[1]。

1.3 个性化治疗方案的智能生成 AI技术在治疗决策中的实践正从标准化模式转向动态个性化配置。通过融合基因组特征、病史及生活方式等多维度数据,智能系统可生成精准化治疗策略。基于机器学习的AI药物推荐系统可提高药物选择的精确性,降低发生不良反应的风险^[8]。在肿瘤治疗领域,基于患者化疗敏感性数据的AI模型可个体化动态调整剂量和化疗方案^[9]。康复医学中,AI结合实时功能监测数据可优化训练计划,但强调医师需审核AI建议,避免技术主导决策^[10]。然而,AI

生成的复杂治疗建议常超越患者认知边界,致使知情同意机制陷入实效性困境^[11]。

1.4 医疗资源分配的智能决策优化 AI技术依托数据驱动范式优化医疗资源配置决策,但其公平性风险亟待审视。基于动态预测框架的资源需求分析系平台可帮助管理者制定优先级策略^[12]。智能分诊模型通过整合患者社会经济数据、既往病史及实时监测数据优化资源配置,缩短平均候诊时间^[13]。然而医疗AI算法因训练数据本身存在的偏差,可能导致资源分配不均,影响公平性^[14]。为确保公平性,应建立算法透明度机制和定期审计制度。

2 法律责任与伦理挑战

2.1 法律界定与责任分配机制的分级重构 AI医疗决策系统的法律主体资格界定仍存争议。若将其定义为辅助工具,责任主体需覆盖开发者、操作者及医疗机构;若视为独立决策主体,则须创设包含技术能力认证的法律框架^[15]。欧盟《人工智能法案》基于人权保护原则将医疗AI列为“高风险系统”,要求开发者承担算法透明度义务并负主要责任^[16];欧洲议会和理事会关于缺陷产品责任的指令提案建议采用无过错责任原则,即生产者应对其产品缺陷造成的损害负责^[17]。我国现行《医疗器械监督管理条例》将AI辅助诊断纳入第三类器械管理,但《人工智能医用软件产品分类界定指导原则》尚未明确动态诊疗场景中的责任主体边界^[18]。因此,建议依据技术自主性程度建立分级追责体系,辅助型系统由操作医师承担最终责任,决策型系统则要求开发者提供可验证算法并承担连带责任^[19]。

2.2 系统致损责任认定与法律规范细化 AI医疗系统因算法缺陷或数据偏差引发的损害事件,需结合技术特征优化责任界定机制。依据我国《产品质量法》与《医疗事故处理条例》,开发者负有确保AI算法合规的法定义务(如数据完整性验证),操作者需履行审慎复核职责。欧盟《人工智能法案》禁止社会评分、无目的生物识别等危害性AI应用,并对高风险AI系统实施全链条责任机制,开发者与部署者未履职时需承担连带责任;数字医疗产品中的AI系统同时也受欧盟《医疗器械法规》和《体外诊断医疗器械法规》监管,且需要遵守《通用数据保护条例》,制造商有义务对技术文档、可追溯性、透明度和人工监督负责^[20]。我国可借鉴

其风险分析级思路, 在医疗 AI 领域适用“过错推定”原则^[21]: 开发者需举证算法符合国家标准 (如 GB/T 25000.51), 并提交第三方认证; 临床操作者需证明其遵循“双人核验”交叉验证流程^[19]。同时建议设立 AI 医疗技术缺陷第三方鉴定机构, 完善损害因果关系认定标准, 并依据《卫生健康行业人工智能应用场景参考指引》^[22]建立医产风险补偿专项基金, 以应对算法不可解释性导致的医疗损害。

2.3 伦理透明度问题与职业生态重塑 智能诊疗系统的伦理困境突出表现为透明度不足与患者自主权受损。有研究探讨了临终关怀环境中患者对人类医生和 AI 所做决策的态度, 结果显示与人类医生相比, AI 提出的终止生命支持的决策更难以获得患者家属认可, 其中部分原因是 AI 能力信任不足和可解释性有限^[23]。破解该困局需立法强制高风险 AI 披露关键算法逻辑, 建立独立审计机制监督透明度执行情况, 并实施动态伦理审查管理^[24]。在职业发展方面, 一项对糖尿病护理使用 AI 的调查显示, 医护人员对 AI 的态度总体上是积极的, 但仍有 34.2% 的受访者对失业表示担忧^[25]。应对路径须通过继续教育强化医务人员 AI 伦理认知, 增设伦理必修课程, 提升对医疗 AI 风险的识别能力^[26]; 同时落实《医疗机构临床决策支持系统应用管理规范 (试行)》要求, 明确 AI 仅作为辅助工具, 医生承担最终决策责任, 并建立多学科参与的 AI 应用审核委员会^[27]。

2.4 资源分配算法的公平性规制 算法并非绝对的“中立”, 有时其带有的歧视性会导致医疗 AI 在诊断和治疗疾病时造成公平失衡现象^[28]。此类算法偏倚急需法律强制介入调整: 首要措施是建立公平性审查制度, 强制公立医院公开资源分配模型核心参数 (如优先级权重设定), 并接受省级伦理委员会年度审计; 其次, 推动 AI 辅助诊断技术通过紧密型县域医共体下沉重点提升偏远乡镇卫生院服务能力, 明确要求县级医院支援乡镇卫生院部署 AI 系统, 以确保基层首诊率达标^[29], 同时辅以财政补贴机制平抑技术应用成本; 此外, 应强制开发者披露数据代表性评估报告 (含城乡、性别、年龄分布), 并将样本均衡性指标纳入算法伦理审查标准, 保障 AI 医疗决策模型遵循医疗公平的伦理基准^[28]。

3 法律与伦理难点及应对策略

3.1 个体化医疗知情同意机制的重构 基于多组学数据的 AI 诊疗系统生成的靶向治疗方案具有技术复杂性, 导致传统知情同意机制失效。对此, 需重构动态知情同意机制: 首先, 可借助可视化交互界面分阶段解析技术风险, 如通过 3D 动画演示药物作用机制, 并赋予患者随时撤回授权的权利; 同步修订《医疗机构管理条例》, 要求 AI 开发者强制披露关键算法逻辑, 并建立独立审计机制监督透明度执行情况^[28]; 此外, 建议在《民法典》中增设“算法透明度披露义务”, 要求高风险 AI 系统提供可解释性报告, 若医疗机构已履行充分风险告知并经患者确认, 可减轻因算法“黑箱”导致的告知责任^[30]。

3.2 算法透明度与伦理审查协同治理 AI 决策的“黑箱”特性显著放大了伦理风险, 应对策略需多维推进。按照欧盟《人工智能法案》建议, 强制高风险 AI 系统提供基本可解释性说明^[20], 确保医生能理解 AI 推荐逻辑; 推动医疗机构组建伦理审查小组, 对 AI 系统进行定期评估; 同时按照《医疗机构临床决策支持系统应用管理规范 (试行)》^[27]要求, 对高风险 AI 医疗产品实施备案管理, 重点核查系统安全性和有效性。

3.3 法律规范体系的适应性迭代 现行法律框架难以匹配 AI 技术的快速迭代节奏, 尚未明确生成式 AI 在动态诊疗中的责任主体, 导致 AI 相关致损案件面临法律适用困境^[31]。建议分阶段实施相应的策略: 短期阶段, 修订《医疗事故处理条例》, 增设“算法缺陷致损”责任条款, 明确开发者、操作方与医疗机构三方责任比例; 中期阶段, 制定“AI 医疗应用管理法”, 按“辅助型/决策型”分级监管, 高风险 AI 需符合 ISO 31000 认证并提交伦理审查报告; 长期阶段, 推动全球 AI 伦理治理标准和数据合规互认, 积极制定与协调国际法规体系。

3.4 医疗资源分配算法的公平性矫正 AI 资源调度模型可能因数据偏倚加剧医疗资源分配失衡。通过构建“动态去偏”技术框架, 实时监测城乡诊疗需求差异, 自动校准资源分配权重^[32]。医疗机构应建立“偏见影响评估”制度, 重点审查算法对弱势群体的服务可及性, 并定期发布公平性审计报告。

- 告^[33]。应立法规定AI模型训练数据集必须满足“最小代表性阈值”要求^[34]，并将健康公平性指标纳入算法性能评估体系。同时应建立“算法偏见举报-复核”机制^[28]，允许对明显失衡的资源分配结果提出异议，由监管部门组织第三方技术验证。最终形成全链条公平性保障体系，从而避免技术应用固化既有医疗资源差距。

4 结 语

AI在医疗决策中的应用通过辅助诊断、治疗方案优化及资源调度等多维度实践，已展现出提升诊疗精准度与效率的技术价值。然而，医疗AI亦引发了法律权责模糊与伦理治理失序的双重困境。未来应构建“法律-伦理-技术”协同框架，通过动态分级界定开发者、操作方与医疗机构权责，依托算法可解释性增强AI医疗决策的透明度，并以适应性立法弥合技术迭代与监管滞后鸿沟。同时应强化跨学科协作与公平性审查，遏制AI算法偏倚导致的资源分配失衡。唯有平衡创新价值与患者权益，方能推动AI医疗决策规范化与可持续发展。

参考文献

 - [1] ZHANG J, ZHANG Z M. Ethics and governance of trustworthy medical artificial intelligence[J]. BMC Med Inform Decis Mak, 2023, 23(1): 7. DOI: 10.1186/s12911-023-02103-9.
 - [2] ALI AMER A Y, WOUTERS F, VRANKEN J, et al. Vital signs prediction and early warning score calculation based on continuous monitoring of hospitalised patients using wearable technology[J]. Sensors (Basel), 2020, 20(22): 6593. DOI: 10.3390/s20226593.
 - [3] SAVOVA G K, DANCIU I, ALAMUDUN F, et al. Use of natural language processing to extract clinical cancer phenotypes from electronic medical records[J]. Cancer Res, 2019, 79(21): 5463-5470. DOI: 10.1158/0008-5472.CAN-19-0579.
 - [4] 秦伟峰, 吕旭东, 段会龙, 等. 药物基因组学临床决策支持系统综述[J]. 中国生物医学工程学报, 2019, 38(1): 102-111. DOI: 10.3969/j.issn.0258-8021.2019.01.013.
 - [5] LIAO S, MO Z, ZENG M, et al. Fast and low-dose medical imaging generation empowered by hybrid deep-learning and iterative reconstruction[J]. Cell Rep Med, 2023, 4(7): 101119. DOI: 10.1016/j.xcrm.2023.101119.
 - [6] BERA K, SCHALPER K A, RIMM D L, et al. Artificial intelligence in digital pathology: new tools for diagnosis and precision oncology[J]. Nat Rev Clin Oncol, 2019, 16(11): 703-715. DOI: 10.1038/s41571-019-0252-y.
 - [7] MENG X, YAN X, ZHANG K, et al. The application of large language models in medicine: a scoping review[J]. iScience, 2024, 27(5): 109713. DOI: 10.1016/j.isci.2024.109713.
 - [8] AHMAD F. Optimizing treatment: the role of pharmacology, genomics, and AI in improving patient outcomes[J]. Drug Dev Res, 2025, 86(3): e70093. DOI: 10.1002/ddr.70093.
 - [9] PADMANABHAN R, MESKIN N, HADDAD W M. Reinforcement learning-based control of drug dosing for cancer chemotherapy treatment[J]. Math Biosci, 2017, 293: 11-20. DOI: 10.1016/j.mbs.2017.08.004.
 - [10] HSU W L, CHEN Y C, LIN C H. AI-based balance rehabilitation system for elderly: personalized training and fall prevention[J]. J Neuroeng Rehabil, 2024, 21(1): 45. DOI: 10.1186/s12984-024-01345-6.
 - [11] CHAU M, RAHMAN M G, DEBNATH T. From black box to clarity: strategies for effective AI informed consent in healthcare[J]. Artif Intell Med, 2025, 167: 103169. DOI: 10.1016/j.artmed.2025.103169.
 - [12] KÖKCIYAN N, SASSOON I, SKLAR E, et al. Applying metalevel argumentation frameworks to support medical decision making[J]. IEEE Intell Syst, 2021, 36(2): 64-71. DOI: 10.1109/MIS.2021.3051420.
 - [13] BANERJEE A, BAIBHAV P, KUMAR N, et al. AI-driven symptom checker: leveraging big data technologies for real-time healthcare triage and diagnostics[C]//2024 International Conference on Sustainable Communication Networks and Application (ICSCNA). December 11-13, 2024, Theni, India. IEEE, 2024: 1289-1293. DOI: 10.1109/ICSCNA63714.2024.10864063.
 - [14] OBERMEYER Z, POWERS B, VOGELI C, et al. Dissecting racial bias in an algorithm used to manage the health of populations[J]. Science, 2019, 366(6464): 447-453. DOI: 10.1126/science.aax2342.
 - [15] 沈宇超. 人工智能在医疗领域应用中的法律问题析[J]. 锦州医科大学学报(社会科学版), 2020, 18(4): 24-30. DOI: 10.13847/j.cnki.lnmu(sse).2020.04.007.
 - [16] SCHUETT J. Risk management in the artificial intelligence act[J]. Eur J Risk Regul, 2024, 15(2): 367-385. DOI: 10.1017/err.2023.1.
 - [17] European Commission. Proposal for a directive of the European Parliament and of the Council on liability for defective products: COM(2022) 495 final[Z/OL]. (2022-09-28)[2025-03-01]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52022PC0495&qid=1751505262867>.
 - [18] 国家药品监督管理局. 国家药监局关于发布人工智能医用软件产品分类界定指导原则的通告(2021年第47号)[EB/OL]. (2021-07-01)[2025-03-01]. <https://www.nmpa.gov.cn/xxfw/xxfw/20210701145947000.html>.

- www.cn-healthcare.com/article/20210708/content-557003.html.
- [19] 胡巧莉. 人工智能服务提供者侵权责任要件的类型构造: 以风险区分为视角[J]. 比较法研究, 2024(6): 57-71.
- [20] ABOY M, MINSEN T, VAYENA E. Navigating the EU AI act: implications for regulated digital medical products[J]. NPJ Digit Med, 2024, 7(1): 237. DOI: 10.1038/s41746-024-01232-3.
- [21] MAROUDAS V P. Fault-based liability for medical malpractice in the age of artificial intelligence: a comparative analysis of German and Greek medical liability law in view of the challenges posed by AI systems[J]. Rev Eur Comp Law, 2024, 57(2). DOI: 10.31743/recl.17223.
- [22] 中华人民共和国国家卫生健康委员会. 国家卫生健康委员会办公厅关于印发卫生健康行业人工智能应用场景参考指引的通知[EB/OL]. (2024-11-14)[2025-03-01]. <https://www.nhc.gov.cn/wjw/c100175/202411/5bcb3c4edd064e31ac5d279caf5830f4.shtml>.
- [23] LAAKASUO M, KUNNARI A, FRANCIS K, et al. Moral psychological exploration of the asymmetry effect in AI-assisted euthanasia decisions[J]. Cognition, 2025, 262: 106177. DOI: 10.1016/j.cognition.2025.106177.
- [24] 宋冠澎, 王启帆. 医疗人工智能算法决策的伦理风险及规制策略[J]. 中国医学伦理学, 2024, 37(9): 1080-1086. DOI: 10.12026/j.issn.1001-8565.2024.09.10.
- [25] MO Y, ZHAO F, YUAN L, et al. Healthcare providers' perceptions of artificial intelligence in diabetes care: a cross-sectional study in China[J]. Int J Nurs Sci, 2025, 12(3): 218-224. DOI: 10.1016/j.ijnss.2025.04.013.
- [26] 赵丹娜, 赵蓝蓝, 陈任, 等. 某三级公立医院医务人员对人工智能医学应用伦理问题的认知研究[J]. 南京医科大学学报(社会科学版), 2022, 22(5): 465-470. DOI: 10.7655/NYDXBSS20220507.
- [27] 中华人民共和国国家卫生健康委员会. 医疗机构临床决策支持系统应用管理规范(试行)[J]. 中国卫生资源, 2023, 26(5): 620. DOI: 10.3969/j.issn.1007-953X.2023.05.032.
- [28] 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. DOI: 10.12014/j.issn.1002-0772.2023.11.15.
- [29] 国家卫生健康委员会. 关于进一步健全机制推动城市医疗资源向县级医院和城乡基层下沉的通知[EB/OL]. (2024-06-03)[2025-03-01]. <https://www.nhc.gov.cn/wjw/c100375/202406/b33eebe413fe4198a4d5b480193fbb97.shtml>.
- [30] 周吉银, 李红英, 杨阳. 人工智能医疗器械的伦理审查要点[J]. 医学与哲学, 2020, 41(6): 35-39, 56. DOI: 10.12014/j.issn.1002-0772.2020.06.10.
- [31] 刘丹, 周吉银. 人工智能医疗器械伦理挑战的对策[J]. 中国医学伦理学, 2020, 33(7): 863-867, 872. DOI: 10.12026/j.issn.1001-8565.2020.07.15.
- [32] 仝媛媛, 郑婉婷, 黄珊蓉, 等. 健康公平: 医疗人工智能中的偏见与治理[J]. 医学与哲学, 2024, 45(7): 36-41. DOI: 10.12014/j.issn.1002-0772.2024.07.08.
- [33] 孙琦, 吕晗, 王振常. 消除健康数据中的算法偏见: STANDING Together 共识的启示[J]. 数字医学与健康, 2025, 3(1): 10-11. DOI: 10.3760/cma.j.cn101909-20241231-00240.
- [34] 陈龙, 曾凯, 李莎, 等. 人工智能算法偏见与健康不公平的成因与对策分析[J]. 中国全科医学, 2023, 26(19): 2423-2427. DOI: 10.12114/j.issn.1007-9572.2023.0007.

[本文编辑] 孙 岩