首都医科大学附属北京天坛医院党委书记岳小林：

AI 赋能医学创新，伦理治理须先行

岳小林

人工智能（AI）正在驱动全球药物研发范式向以人为中心的核心导向转型。传统药物研发普遍面临研发周期长、成本高、成功率低以及动物与人存在物种差异等瓶颈，而 AI、大数据、类器官、器官芯片、数字孪生等新技术的发展，为药物研发提供了新的路径。

从技术赋能的具体维度来看，临床队列、多组学数据和 AI 驱动靶点发现，可大幅提升药物靶点筛选效率；类器官与器官芯片能够更加真实地模拟人体组织功能，有效弥补动物实验的缺陷；数字孪生和虚拟临床试验则有望实现药物疗效和安全性的预测，提高研发成功率，推动药物研发向精准化、高效化、伦理友好化的方向发展。

AI 是一把“双刃剑”，在提升医疗效率和科研能力的同时，也带来了数据安全、患者隐私、算法偏倚、模型可靠性、责任归属等一系列新的伦理挑战。

医疗 AI 高度依赖海量医疗数据，数据质量直接影响算法性能，而数据采集、共享和应用过程中涉及患者知情同意、个人隐私保护和信息安全等问题，已经成为医学伦理审查的重要内容。特别是基于历史医疗数据开展的回顾性研究和人工智能模型训练，如何规范泛知情同意和免除知情同意的适用范围，如何依法合规开展数据共享与利用，都对现有伦理审查制度提出了新的要求。

在医疗 AI 快速发展的背景下，伦理审查的重点正从传统的研究参与者保护，进一步延伸到数据治理和 AI 治理。医疗机构应建立覆盖数据采集、提取、脱敏、传输、存储、使用和销毁全流程的规范化管理机制，数据管理应由专业部门统一负责，确保研究数据在全生命周期内安全、规范、可追溯。同时，针对可能产生重大社会影响的 AI 研究和应用，应进一步完善科技伦理审查机制，加强风险识别与动态监管，在保障创新活力的同时守住伦理底线。

AI 的发展正在深刻改变医学科研和医疗服务模式，但技术创新必须始终坚持“伦理先行”。未来，应进一步完善 AI 相关法律法规、伦理规范和行业标准，加强科技、卫生、药监、数据、法律、伦理等多部门协同治理，推动 AI 在合法合规、可信可控的前提下赋能医学研究和临床实践，更好地服务于全民健康的核心需求。

北京协和医学院卫生法与生命伦理学系主任睢素利：

使用 AI 兼顾安全与自主

睢素利

随着人工智能（AI）技术的飞速演进，生成式人工智能（Gen AI）在医疗健康领域展现出显著的应用价值，从辅助诊疗到优化文书流程，技术正深度赋能医疗行业。然而，伴随技术红利而来的还有隐私泄露、算法偏见、技术“黑箱”及“AI 幻觉”等严峻的伦理与治理难题。

AI 在医疗健康领域具有广阔的应用前景。随着技术的飞速发展，引入 AI 可以显著提高工作效率并改善医疗环境。它不仅应用于临床医疗，还覆盖伦理审查、药物开发等方面。

作为 AI 领域的新兴成果，Gen AI 具备文本、图片、音频、视频的生成能力。其核心并非“检索答案”，而是基于概率生成新的内容。

Gen AI 在医疗中的应用可以显著减轻医务人员的负担。它可以优化电子病历流程、自动回复患者咨询、翻译病历材料以及辅助数据检索，从而让医生从繁琐的文书工作和数据处理中解脱出来。

当前，我国 AI 医疗正进入政策驱动的快速发展期。2025 年 8 月，国务院印发《关于深入实施“人工智能+”行动的意见》，要求探索推广人人可享的高水平居民健康助手，有序推动 AI 在辅助诊疗、健康管理等场景的应用。

然而，在政策扶持加速技术落地的背景下，监管与制度体系建设相对滞后。当技术能力的扩展速度显著快于规范与制度建设时，科技创新本身反而可能成为社会风险的放大器。

我们需要关注技术发展初期的潜在问题，避免陷入“科林格里奇困境（Collingridge Dilemma）”。这种困境是指当技术还处于初期时，其社会影响难以预测；而当技术已经充分发展并产生严重问题时，对其进行控制和治理的成本将变得非常巨大，甚至难以管控。因此，我们必须前瞻性地考虑风险。

首先是隐私泄露的风险。生物医学大数据是训练 Gen AI 算法和验证算法模型的基础。医用 Gen AI 涉及大量个人健康信息，包括病史、诊断结果、治疗计划等。在数据的收集、使用、分析和共享过程中，如果保护工作不到位，个人隐私将面临巨大的泄露风险。

其次是偏见与歧视的问题。如果原始训练数据本身存在偏差（如代表性、均衡性或多样性不足），算法就会将这些偏见带入决策过程中，导致系统性的偏见。此外，Gen AI 的设计者作为出资方或技术方，其个人的主观倾向、利益诉求或潜意识中的倾向，也可能通过算法设计影响输出结果。

“AI 幻觉”与“谄媚”现象是 Gen AI 特有的风险。所谓“幻觉”，是指模型会一本正经地编造虚假信息，例如生成根本不存在的法律条文或学术参考文献。而“谄媚”现象则是指模型在交互中过度取悦用户，迎合用户的语言倾向。在医疗场景下，如果 Gen AI 为了顺从患者而生成其“喜欢”但错误的内容，可能会导致患者产生过度依赖，甚至造成严重的医疗后果。

黑箱算法导致的可解释性缺失，挑战了“知情同意”原则。如果医生无法理解 Gen AI 给出建议的逻辑，就难以进行实质性的复合审查。这不仅可能损害患者的知情同意权，还可能削弱人类的自主决策能力。

Gen AI 的应用还引发了责任缺口的问题。随着技术深度融入决策，传统的“医生－医院”二元责任链条可能会断裂。当发生医疗纠纷时，责任主体往往变得模糊，医院、医生与技术提供方之间的界限不清，导致责任难以追溯与界定。

在制定 Gen AI 应用规则时，必须确立核心原则。我们应优先保护人类自主性，促进人类福祉、安全和公共利益，并确保算法的透明度、可解释性，建立问责机制。

（3~6 版由本报编辑部采访整理）