

• 特别专题 •

快速发展中的现代临床研究管理

(三)数据、数字化与人工智能带来的机遇、挑战与对研究管理的启示

武阳丰

北京大学临床医学高等研究院,北京 100191

通信作者:武阳丰,Email: wuyf@bjmu.edu.cn,电话:010-82805836

【摘要】 临床研究已迈入一个健康数据空前增长、数字技术飞速进步及人工智能快速发展的新时代。然而,更多的数据并不会自动产生好证据。复杂的算法也不一定产生可靠、公平或临床有用的结果。本文探讨数据、数字化与人工智能在临床研究中所扮演的角色及其演变,发现数字化转型贯穿整个研究生命周期,从根本上改变了临床研究的范式,创造了新的机遇。临床研究的核心挑战不再是数据生成,而是数据治理。在大数据和人工智能时代,成功的临床研究不仅需要更智能的算法,还需要更智能的治理。

【关键词】 临床研究; 科研管理; 数据; 数字化; 人工智能

【中图分类号】 R19;R-05 **【文献标识码】** A DOI:10.3760/cma.j.cn113565-20260610-00193

Modern clinical research management in a rapidly changing world: (Ⅲ) Data, digital transformation, and artificial intelligence in clinical research: opportunities, challenges, and implications for research management

Yangfeng Wu, The Institute of Advanced Clinical Medicine, Peking University, Beijing 100191, China

Corresponding author: Wu Yangfeng, Email: wuyf@bjmu.edu.cn, Tel: 0086-10-82805836

【Abstract】 Clinical research has entered a new era characterized by unprecedented growth in health data, rapid advances in digital technologies, and the rapid development of artificial intelligence. However, more data do not automatically produce good evidence, and complex algorithms do not necessarily yield reliable, fair, or clinically useful results. This article examines the evolving roles of data, digital transformation, and artificial intelligence in clinical research, finding that digital transformation permeates the entire research lifecycle, fundamentally reshapes the paradigm of clinical research, and creates new opportunities. The core challenge of clinical research is no longer data generation, but data governance. In the era of big data and artificial intelligence, successful clinical research requires not only smarter algorithms, but also smarter governance.

【Key words】 Clinical research; Research management; Data; Digital transformation; Artificial Intelligence
DOI:10.3760/cma.j.cn113565-20260610-00193

临床研究已迈入一个健康数据空前增长、数字技术飞速进步及人工智能快速发展的新时代。电子健康记录(EHR)、疾病登记平台、基因组数据库、可穿戴设备和移动健康应用等正以前所未有的规模生成海量数据。与此同时,计算力、机器学习和数据科学领域的创新,正深刻改变着研究者采集、分析和解读数据的方式方法。这些进步正在重塑临床研发生态。

未来前景广阔。大规模健康数据不仅能够支持生成真实世界证据,而且可以获得传统临床试验难以触及的患者群体的健康信息。数字技术在加速证据生成、降低研究成本、提高参与者参与度、提升研究效率和促进个体化医疗等方面有巨大潜力。人工智能为疾病预测、诊断、治疗选择及新药开发和试验方案优化提供了新方法。

然而,更多的数据并不会自动变成更好的证据。复杂的算法也不一定产生可靠、公平或临床有用的结

果。数据碎片化、质量堪忧、隐私风险、监管不确定性及人才队伍不足,都可能使技术创新带来的好处大打折扣。为此,本文试图探讨临床研究中数字化转型与人工智能带来的机遇、挑战与对研究管理的启示。

1 临床研究的数字化转型

1.1 数据从稀缺到充裕

在过去大部分时间里,研究者通常需要花费数年时间,通过精心设计的研究来招募参与者和收集信息。这个过程费钱、费时、费力,因此研究人群规模通常较小。

如今,医疗卫生系统通过 EHR、实验室系统、影像平台、药房数据库、健康保险理赔数据、疾病监测系统、基因测序平台及行政记录,常规化生成海量数据。在医疗机构之外,可穿戴传感器、智能手机及数字健康应用持续捕捉着人们的生理参数、行为和环境暴露等信息。过去需要数年时间收集的数据,现在数周或数月即可获得。

1.2 数字化贯穿整个研究生命周期

数字化转型影响着临床研究的各个阶段。在规划阶段, EHR 可用于评估可行性、识别合格人群和估计招募潜力。在实施阶段, 数字平台有助于参与者招募、电子知情同意、远程监查及实时数据采集。在分析阶段, 先进的计算方法使研究者能够整合多个数据源并识别复杂模式。研究完成后, 数字化传播渠道有助于快速分享研究成果并扩大与利益相关方的互动。

1.3 学习型医疗系统的兴起

EHR 是临床研究数据最重要的来源之一。管理得当的 EHR 系统可支持观察性研究、效果比较研究、实效性临床试验、上市后监测及质量改进计划等, 也推动了“学习型医疗系统”概念的出现, 即临床诊疗与研究日益融合。在学习型医疗系统中, 常规诊疗中产生的数据不断为研究提供信息, 而研究又迅速影响临床实践。这类系统有望加速证据生成, 缩小科学发现与医疗实践之间的差距。

数字化转型从根本上改变了临床研究的研究范式, 为临床研究创造了新的机遇。研究者现在能够接触到更大规模和更多样化的人群; 过去需要专门前瞻性研究的问题, 现在可利用现有数据集进行探索; 通过医疗记录关联, 还可实现长期随访, 开展更长周期的纵向分析; 通过跨机构和跨地区汇总信息的大规模数据库, 可研究罕见结局和少见疾病。数字化转型的潜力在于它能使临床研究更快、更高效、更能代表真实世界实践。

2 数据治理——数字化研究的基石

2.1 治理比体量更重要

那种“数据量越大, 证据质量越高”的想法并不成立。数据完整性、准确性、一致性和标准化, 仍然是决定研究结果真实性的关键因素。在许多情况下, 经过精心设计的小型研究得出的结论比质量低劣的大数据集产生的结论更加可靠。

大型数据集常常包含缺失值、不一致的编码、测量误差和系统性偏倚。为临床诊疗或行政目的收集的数据可能与研究目标不符。不同机构、医疗体系和国家之间的差异会进一步使数据整合复杂化。

因此, 有效的数据治理是现代临床研究管理中最重要职责之一。缺乏治理, 大规模数据项目可能会失去公众信任和科学可信度。

2.2 数据治理的关键维度

传统上, 数据质量常被视为统计人员或数据管理员承担的技术责任。实际上, 数据质量从根本上说是一个治理问题。数据治理主要包括数据质量管理、互操作性和数据的访问与共享 3 个方面。

2.2.1 数据质量管理 数据质量必须主动管理, 而非事后评估。有效的方法包括标准化的数据收集流程、清晰的定义和术语体系、数据验证和清洗方案、质量保证体系、员工培训和持续监测等。

研究管理者在机构内建立数据质量文化方面发挥着关键作用。对基础设施、人员和质量保证体系的投入, 往往比投资于日益复杂的分析工具产生更大的长期效益。

2.2.2 互操作性 医疗数据常常分散在不同机构、不同部门和不同软件系统中, 使用不同的格式和术语。一位患者的信息可能分布于医院病历、基层医疗记录、实验室系统、影像档案、药房数据库和公共卫生登记库。没有互操作性, 这些数据将保持孤立, 无法充分利用, 成为现代临床研究的主要障碍之一。

提高互操作性需要标准化的编码系统、通用数据模型、统一的元数据结构和安全的数据关联机制。未来临床研究系统的成功在较大程度上取决于整合不同数据源的能力。这一挑战主要是管理和组织层面的, 而非技术层面的。

2.2.3 数据访问与共享 数字研究中的一个核心问题是在数据科学价值最大化与患者隐私保护之间取得平衡。过度限制可能阻碍科学进步, 而保障不足则可能损害公众信任。

有效的治理应针对数据所有权、访问程序、数据使用协议、问责机制和数据的二次使用等建立透明的规则, 强调基于风险的方法、数据最小化原则、安全的研究环境、受控访问模式和数据使用透明度, 使数据能够为了科学利益负责任地共享。

3 人工智能与临床研究

人工智能已成为医疗和生物医学研究中最具影响力的技术进步之一。围绕人工智能的热情, 反映了其在识别传统统计方法难以发现的复杂关系方面的潜力。然而, 人工智能不仅仅是一种新的分析工具。它引入了根本上不同的证据生成和决策方法, 引发了重要的科学、伦理和管理问题。

3.1 人工智能的挑战

3.1.1 预测与解释 区别预测与解释之间的差异在人工智能与临床研究中很重要, 因为临床研究不仅是结局预测研究, 更需要理解因果关系。一个高度准确的预测模型可能对疾病生物学机制的揭示作用甚微。反之, 一个具有临床意义的因果关系可能并不产生强大的预测效能。因此, 研究管理者和研究者须避免将算法准确性与科学有效性等同起来。

3.1.2 算法偏倚 人工智能系统从历史数据中学

习。如果这些数据包含偏倚,那么产生的算法可能会延续甚至放大潜在的偏倚。不平等的医疗可及性、特定人群代表性不足、历史上的治疗差异、测量误差等都可能带来偏倚。在临床研究中,这种偏倚有可能导致或扩大健康不公平。

3.1.3 黑箱挑战 许多现代人工智能模型难以解读。这种“黑箱”问题给临床医生、患者、监管机构和伦理委员会带来挑战。临床采用通常不仅需要对其性能的信心,还需要对算法的理解。

3.1.4 可重复性 专有算法、报告不完整、文档不充分、代码不可得和不透明的模型开发过程等会使生物医学研究中本已存在的可重复性问题被人工智能进一步放大。许多模型在用于开发的数据集上表现极佳,但应用到新的人群或场景时,性能便会下降。这是因为算法常常学到的是反映局部特征的模式,而非具有普遍性的生物学关系。

3.2 人工智能的监管

3.2.1 隐私与公众信任 人工智能时代,隐私保护已成为一个核心关切。即使是去标识化的数据集,当与其他信息源合并时,也可能存在再识别的风险。公众信任取决于能否证明数据是安全存储的、负责任地使用的、透明治理的和没有滥用的。

3.2.2 创新与安全 传统的临床研究监管方法是为相对静态的干预措施设计的。人工智能系统带来了新的挑战,因为它们可能持续学习,随时间更新,依赖于变化的数据源。监管的关键问题包括:应如何监管具有自适应能力的算法?基于人工智能的临床工具需要哪些证据?应如何管理软件更新?谁对算法的决策负责?监管机构必须在创新与患者安全之间取得平衡。

3.2.3 透明与规范 越来越多的期刊、监管机构和资助机构正在强调人工智能研究的透明报告标准,如 TRIPOD-LLM 等。机构的研究管理者应加强培训并制定内部政策支持透明报告。

3.2.4 偏倚与公平 处理偏倚需建立多样化的数据集开发、聚焦公平性的验证、跨亚组的性能监测、独立稽查、实施后的持续评估等方面的管理程序。健康公平性也应被视为与准确性、可靠性和效率同等重要的质量指标。

3.2.5 内部与外部验证 可靠的验证需独立的测试数据集、跨机构的外部验证、在不同人群中的评估及在真实世界环境中进行前瞻性评估。研究管理者

应要求在将人工智能系统用于临床环境前进行严格的验证。

4 数字时代的人才队伍转型

4.1 现代临床研究所需的新技能

临床研究的数字化转型正在重塑人才需求。未来的研究团队日益需要以下领域的专业知识:

- 数据科学
- 健康信息学
- 机器学习
- 数据治理
- 网络安全
- 监管科学

随着研究变得更具数据密集性和技术复杂性,传统的学科界限变得不再那么重要。

4.2 建设跨学科团队

没有哪位研究者能够掌握现代临床研究的所有方面。成功越来越依赖于多学科协作,参与者包括临床医生、流行病学家、生物统计学家、数据科学家、软件工程师、伦理专家及监管专家等。研究管理必须创建促进协作的结构,同时支持不同专业群体的职业发展。

4.3 数字时代的领导力

单靠技术创新无法改变临床研究。领导力对于设定战略优先事项、分配资源、建设基础设施、建立治理体系及促进负责任创新是必需的。最成功的组织将是那些认识到数字化转型既是管理挑战也是技术挑战的组织。未来的领导者不仅要懂科学和医学,还要懂数据治理、组织变革和系统思维。

5 结论

数据、数字技术和人工智能正在以前所未有的速度改变临床研究。大规模健康数据、先进分析和机器学习为加速发现、改善预测、提高试验效率和支持精准医学提供了前所未有的机遇。

然而,这些创新的价值从根本上取决于底层数据的质量、治理体系的稳健性及机构管理的有效性。大数据不会自动产生好证据,强大的算法也不会自动生成可信赖的知识。数据治理必须被视为战略基础条件,而非行政辅助支持。

临床研究的未来将不仅由其技术的先进性决定,更取决于研究系统负责任、公平且有效地管理这些技术的能力。在大数据和人工智能时代,成功的临床研究不仅需要更智能的算法,还需要更智能的治理。

(收稿日期:2026-06-10)