

· 述 评 ·

大数据时代的临床医学研究——机遇和挑战

沈洪兵*

南京医科大学公共卫生学院流行病学系,江苏 南京 211166

[摘要] 临床医学研究是医学发展的核心驱动力。在大数据背景下,国家高度重视、重点支持,通过顶层设计为我国临床医学研究提供制度保障和资源支撑。多组学与信息技术的飞速发展,为我国临床医学研究提供了宝贵的资源,助力临床医学研究内容的持续创新,催化临床诊疗模式的不断变革,我国临床医学研究正面临新的机遇。然而,目前我国生物医学大数据还存在存储碎片化、数据交汇机制缺乏等问题,亟需建立政府主导的数据共享机制,寻求通力合作的研究模式。在新的时代背景下,依托大数据的真实世界研究将逐渐成为临床干预的重要证据来源,我们要加强临床数据资源和生物样本资源的积累,创新研究方法,打破传统思维模式,迎接新的机遇与挑战。

[关键词] 生物医学大数据;临床医学研究;资源积累;数据共享

[中图分类号] R-05

[文献标志码] A

[文章编号] 1007-4368(2020)03-303-03

doi: 10.7655/NYDXBNS20200301

Clinical medical research in the age of big data—opportunities and challenges

SHEN Hongbing*

Department of Epidemiology, School of Public Health, Nanjing Medical University, Nanjing 211166, China

[Abstract] Clinical medical research is the core driving force of medical development. In the context of big data, the China government provided institutional guarantee and resource support for the clinical medical research through top-level design. The rapid development of multi-omics and information technology has provided valuable resource accumulation, helped the continuous innovation, and catalyzed the continuous change for China's clinical medical research. China's clinical medical research is facing new opportunities. However, China's biomedical big data still has challenges such as fragmentation of storage and lack of data exchange mechanisms. It is urgent to establish a government-led data sharing mechanism and seek a cooperative research model. In the new era, real-world research relying on big data will gradually become an important source of evidence for clinical intervention. We must strengthen the accumulation of clinical data as well as biological sample, innovate research methods, break the traditional thinking mode, and meet new opportunities and challenge.

[Key words] biomedical big data; clinical medical research; resource accumulation; data sharing

[J Nanjing Med Univ, 2020, 40(03): 303-305]

临床医学研究是指以疾病的诊断、治疗、预后和病因为主要研究内容,以患者为主要研究对象,以医疗服务机构为主要研究基地,由多学科人员共同参与组织实施的科学研究活动^[1]。临床医学研究通常包括临床应用型研究和临床指导型研究,前者指临床新技术和新方法,后者指临床新观念和新理念^[2]。开展临床医学研究是提高医疗质量和水平、

[基金项目] 国家自然科学基金创新研究群体(81521004)

*通信作者(Corresponding author), E-mail: hbshen@njmu.edu.cn

拓新临床理念的主要路径,也是临床医学人才培养和学科建设的重要手段。

在生物医学大数据的背景下,国家高度重视、重点支持,依托庞大的患者数量和临床数据资源,我国临床医学研究正面临新的机遇。一方面,临床医学研究是卫生与健康科技创新的重要组成部分,近年来,国家在顶层设计上高度重视、重点投入支持。依托国家重点研发计划,“十三五”期间国家启动了精准医学重点专项、生殖健康与出生缺陷防控

重点专项以及重大慢性病防控重点专项,建立了国家百万自然人群队列、重大疾病专病队列、出生队列以及对应的大数据平台和知识库。这些大型人群队列所建立的生物样本库和数据资源库,将成为疾病的个体化预防、分类、诊断和治疗的基础资源。此外,国家四部委联合发布的《国家临床医学研究中心五年(2017—2022年)发展规划》提出将在2021年底围绕主要疾病领域统筹建成100个国家临床医学研究中心,完善国家医学创新体系,增强临床医学研究创新能力。国家卫健委在《关于全面推进卫生与健康科技创新的指导意见》中要求加强临床医学研究资源的整合与共享,鼓励多学科、多中心合作,努力打造临床研究创新团队,加强临床医学研究队伍建设。这些政策为推动我国临床医学研究快速发展提供了重要的制度保障和资源支撑。

另一方面,近年来,多组学和大数据技术的飞速发展,推动了基因组、转录组、蛋白质组、表观遗传组、代谢组等海量生命组学数据的快速增长。同时,信息技术的革新,将社会活动和医疗活动的过程数字化,医疗卫生服务平台数据、居民健康管理数据、公共卫生普查数据、医院信息系统数据、临床医学研究数据、生物信息数据等各类医疗数据被越来越多的收集和存储。这些数据容量庞大,种类繁多,产生和更新速度快,蕴藏着涉及人类健康的多层次、高维度信息,具有重要的科学价值。我们已经进入了具备相当深度和广度的生物医学大数据时代,也必将极大地推动基于大数据的临床医学研究^[3]。

然而如何将我国的临床医学研究与生物医学大数据和临床医学大数据高效结合,还存在着诸多挑战。庞大的患者数量使我国拥有世界上潜在规模最大的临床数据资源,然而总体上缺乏具备科研标准的临床诊疗系统;不同医院甚至不同科室临床信息系统割裂,质量控制不够严格,错误率高,缺乏结构化的数据展示;患者随访追踪不规范、失访率高;同时,医院来源的临床信息数据缺乏与疾病监测、死亡报告系统及医保系统的链接,造成了目前我国生物医学大数据存在管理分散、存储碎片化、数据交汇机制缺乏、标准化管理混乱、质量参差不齐等问题,限制了我国临床医学大数据资源的开发利用^[3]。面对当前生物医学和临床医学大数据的挑战,亟需建立政府主导的数据共享机制,开发建立生物医学大数据管理、交汇、共享与挖掘的技术与资源体系,寻求数据科学家与临床科学家广泛参

与、通力合作的研究模式。

生物医学大数据的快速发展催化了临床诊疗模式的不断变革,医学进入精准医学时代。精准医学是以大数据的积累分析为基础,以个体化治疗为核心,根据不同患者的遗传背景、环境因素、生活方式、临床表现和组学特征,在对疾病进行重新“分类”的基础上实施“对症用药”,实现“量体裁衣”式的个性化医疗模式^[4]。精准医学更加尊重患者的个性化特征和需求,对一种疾病不同状态和过程进行精确亚分类,最终实现对疾病和特定患者进行个性化治疗的目的。精准医学模式下的临床医学研究加深了人类对疾病的认识和理解,更利于全面评价和比较现有的治疗方案,更好地论证新治疗方法的疗效和安全性,进而优化疾病的防治策略,完善治疗领域的临床实践指南^[1]。

大数据和人工智能的技术进步正在推动临床医学研究的持续创新。以深度神经网络为代表的的人工智能技术,在医学影像处理的应用方面已经呈现巨大的应用潜力。来自天津医科大学肿瘤医院的陈可欣教授的研究团队利用30余万张甲状腺超声图像,基于深度学习算法构建了甲状腺癌的人工智能识别系统,其灵敏度和特异度可以媲美经验丰富的影像专家,实现甲状腺癌人工智能精确诊断,为改善基层医疗机构甲状腺超声诊疗水平提供了一种快速、准确和便利的诊断工具^[5]。类似的成功应用正呈现爆发式增长,已经在致盲性视网膜疾病与肺炎、阿尔茨海默病、皮肤癌、脑膜瘤等医学影像领域取得了突破进展^[6-7]。此外,依托数百万患者的大数据信息,通过深度学习,人工智能系统在计算机模拟药物筛选与研发、健康管理、智慧医疗等领域的潜在价值正在不断显现^[8]。

在当今临床医学大数据的背景下,真实世界研究(real world study, RWS)将成为临床干预的重要证据来源。临床医学研究的发展经历了传统临床研究、循证医学研究和转化医学研究等不同阶段。目前,国际公认大样本随机对照试验(randomized controlled trial, RCT)和据此开展的循证医学研究是评价临床干预有效性和安全性最可靠的依据。然而,RCT的研究样本通常具有高度选择性,依靠这样的小样本研究结果指导大样本人群的临床诊疗,其外部真实性常受质疑。与RCT研究不同,RWS不基于特定的患者群和研究环境,而是依据真实医疗过程,在大样本患者的真实医疗环境下,观察具体治疗干预措施在真实医疗过程中的外部有效性和安

全性。临床医学大数据的发展,为RWS提供了可利用的多源数据,保证了RWS数据的多元性、重要性和时效性。

依托生物医学和临床医学大数据开展临床医学研究,既要加强临床数据资源和生物样本资源的积累,又要遵循合适的研究方法。临床流行病学是临床医学研究的重要工具,是多学科交叉结合的临床科研方法,临床科学家从患者个体的临床诊治入手,将流行病学和统计学的原理和方法应用到临床观察,以探索疾病的病因、诊断、治疗和预后等规律。其核心内容是从临床实践入手确定临床问题,采用科学的设计确定研究对象和信息来源,选定合理的暴露和结局指标进行测量和评价,并在研究实施过程中进行严格的质量控制,确保研究的真实性和可靠性。及时、合理地整合利用临床资源,尤其是加强临床专病队列的建设,系统收集高质量的临床基线数据资料,开展系统流行病学调查,建立合格的生物样本库,积累多维度检验检测数据,进行长期随访,追踪多种临床表型和结局,对于提高临床科研水平和科研效率具有战略性意义。

我们相信随着技术的不断进步,体制机制的不断完善,生物医学和临床医学大数据必将助力我国临床医学研究的快速发展,开辟我国临床医学研究新局面。作为临床科研工作者,我们要打破传统思维模式,应用创新的思考方式,迎接其带来的机遇与挑战。

[参考文献]

- [1] 仇小强. 大数据和精准医学时代临床研究思维的转变[J]. 中国癌症防治杂志, 2017, 9(2): 85-89
- [2] 宁光. 大数据与智慧医疗时代临床研究的策略性思考[J]. 中华内分泌代谢杂志, 2018, 34(7): 537-538
- [3] 张国庆, 李亦学, 王泽峰, 等. 生物医学大数据发展的新挑战与趋势[J]. 中国科学院院刊, 2018, 33(8): 853-859
- [4] National Research Council (US) Committee on A Framework for Developing a New Taxonomy of Disease. Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease [R]. Washington(DC): National Academies Press(US), 2011
- [5] LI X, ZHANG S, ZHANG Q, et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study [J]. Lancet Oncol, 2019, 20(2): 193-201
- [6] ESTEVA A, KUPREL B, NOVOA R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. Nature, 2017, 542(7639): 115-118
- [7] KERMAN Y D S, GOLDBAUM M, CAI W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. Cell, 2018, 172(5): 1122-1131
- [8] 李志勇, 李鹏伟, 高小燕, 等. 人工智能医学技术发展的聚焦领域与趋势分析[J]. 中国医学装备, 2018, 15(7): 136-145

[收稿日期] 2020-02-20

欢迎投稿 欢迎订阅