

真实世界研究与中医药大数据

张冬¹ 张俊华¹ 孙凤² 田金徽³ 庞稳泰¹ 杨丰文¹ 金鑫瑶¹ 刘智¹ 郑文科¹

(1 天津中医药大学循证医学中心,天津,301617; 2 北京大学循证医学中心,北京,100191; 3 兰州大学循证医学中心,兰州,730000)

摘要 数据引领决策成为医药卫生领域的特征,健康医疗大数据逐渐成为热点。基于大数据的真实世界研究也成为医疗卫生行业关注的焦点,因其涉及到药品注册审批、医保支付及健康促进等卫生政策制定。大数据技术与真实世界研究相辅相成、相互促进。基于大数据的真实世界研究为中医药临床评价提供了新的思路和方法,然而还处于起步和探索阶段,需要深化对真实世界研究和大数据的认识。本文系统总结了基于大数据技术的真实世界研究进展以及存在问题,提出了相应的建议与应对策略,为推动基于大数据的真实世界研究提供参考。

关键词 中医药;真实世界研究;真实世界数据;真实世界证据;大数据;数据库;问题与对策;循证医学

Real-World Study and Big Data of Traditional Chinese Medicine

Zhang Dong¹, Zhang Junhua¹, Sun Feng², Tian Jinhui³, Pang Wentai¹, Yang Fengwen¹,
Jin Xinyao¹, Liu Zhi¹, Zheng Wenke¹

(1 Evidence-based Medicine Center, Tianjin University of Traditional Chinese Medicine, Tianjin 301617, China; 2 Evidence-based Medicine Center, Peking University, Beijing 100191, China; 3 Evidence-based Medicine Center, Lanzhou University, Lanzhou 730000, China)

Abstract Data leading decision-making has become a feature of the field of medicine and health, and big data of health care has gradually become a hot topic. Real-world study based on big data has also become a focus of attention in the medical and health industry, because it involves the formulation of health policies such as drug registration approval, medical insurance payments, and health promotion. Big data technology and real-world study are complementary and mutually reinforcing. Real-world study based on big data provides new ideas and methods for clinical evaluation of traditional Chinese medicine. However, it is still in its infancy and exploration stage, and deeper understanding of real-world study and big data is needed. In this paper, we systematically summarized the real-world study progress and existing problems based on big data technology, and proposed corresponding suggestions and coping strategies to provide a reference for promoting real-world study based on big data.

Key Words Traditional Chinese medicine; Real-world study; Real-world data; Real-world evidence; Big data; Database; Problem and solution; Evidence-based medicine

中图分类号:R2-03 文献标识码:A doi:10.3969/j.issn.1673-7202.2019.12.005

信息技术和产品的快速发展,深刻影响着社会生产和生活的方方面面,也影响着医学模式和决策思路的转变。大数据具有容量大、来源广、增速快、价值密度低等特点^[1],支持从大规模数据中挖掘知识,寻找规律^[2]。当前大数据技术发展应用方兴未艾,对医疗大数据价值的认识逐渐深化,学术界、产业界、管理部门都已展开相应的研究和推进工作。如何把医疗大数据转变为可供医疗相关决策可用的证据,需要开展基于大数据的研究,这是真实世界研究的重要内容。本文围绕真实世界研究的特征,探讨中医药大数据研究的现状和问题,并提出建议。

1 真实世界研究与大数据的关系

真实世界研究与传统临床试验的差异根本在于

研究环境是否符合临床和生活的实际情形^[3]。真实世界研究包括多样的设计类型,既有临床试验,也有观察性研究。真实世界研究之所以被广泛重视,是因为医疗大数据的积累为从中发现知识,形成证据,指导决策提供了可能。因此,对真实世界研究的重视,重点是基于大数据技术的发展,为临床评价和健康研究提供了新的路径。

真实世界医疗大数据由于来源多种多样,动态实时变化、且没有标准限制,因此真实世界研究不可避免的遇到了海量数据和非结构化数据无法有效分析的挑战^[4-5]。大数据技术的实际应用,从数据存储、数据处理、数据分析等多个方面为真实世界研究提供了技术支持。同时,随着真实世界研究的不断

基金项目:国家自然科学基金项目(81473544);天津市高等学校“创新团队培养计划”(TD13-5047)

作者简介:张冬(1989.06—),男,硕士,研究实习员,研究方向:中医药评价指标研究,E-mail:zhangd@tjcm.com.cn

通信作者:张俊华(1980.02—),男,博士,研究员,研究方向:循证中医药学,E-mail:zjhtcm@foxmail.com

发展,对大数据技术提出了新的需求,促进了大数据技术的发展与应用。因此,真实世界研究需要以大数据技术为工具,得以高效地从数据中发现问题、总结规律,而大数据技术则以真实世界研究为应用场景,不断丰富技术积累、更新算法框架、推动研究深入。

2 真实世界大数据的来源

医疗领域大数据的来源途径非常广泛、数据类型多样,与过去传统的数据有着很大的区别,涵盖的内容丰富而繁杂。图1以真实世界研究大数据来源单位为例,主要包含医疗从业者提供的电子病历档案、临床评估记录、临床检查结果、个人健康档案等;患者在社交网络的行为数据、可穿戴设备产生的结果数据以及健康体检数据等;医药企业提供的药品研发、营销、流通等各个环节的数据;政府管理机关持有的报销数据以及制定的相关政策等;基因组数据包括个人基因测序数据等^[6-7]。这些复杂、海量的数据并不是相互独立、难以分析的,开展基于大数据的真实世界研究,能够很好的挖掘其中隐藏的价值信息,为临床医疗管理、医保支付决策以及政府卫生政策制定提供证据支持,促进决策机制向科学化转变。

3 中医药大数据平台建设进展

2015年国务院发布了《关于促进和规范健康医疗大数据应用发展的指导意见》,并在“中医药信息化建设‘十三五规划’”中明确提出要建立国家中医药智慧云服务平台,对中医药大数据的发展指出了方向^[8]。

目前,基于大数据的中医药安全性、有效性和系统评价研究在全国科研机构、院校与相关企业开展。中国中医科学院中医药信息研究所进行了中医药数

据库的开发,完成了包括中医药期刊文献、疾病诊疗、中药、方剂等在内的48个子数据库,共120余万条数据,为中医药研究提供了基础数据^[9]。中国中医科学院中医临床基础医学研究所建成了国内首个科研用医疗电子数据仓库,其中收集了400余万份电子病历,并对临床用药的安全性和有效性进行了研究,开展了真实世界用药特征的分析^[10]。中国中医科学院大数据中心组织建设了中医药大数据中心,建立了中医药临床数据共享系统,完成了6000余种临床疾病模板库,形成了中医药数据标准和共享管理,搭建了具有中医特色的健康云平台^[11]。北京大数据研究院成立了中医大数据中心,建设包括中医人工智能、中西药临床服务以及中西药临床科研一体化大数据库及分析平台^[12]。中国中医科学院中药研究所成立了中国中药研究中心,搭建中药数据库和中药基因组对中药各相关领域进行数据挖掘^[13]。国家人口健康科学数据中心依托中国医学科学院医学信息研究所研发建设了人口健康科学数据仓储系统,实现了对人口健康领域科学数据的采集、加工、管理等功能,促进了科学数据的开放共享^[14]。天津中医药大学循证医学中心建立了中医临床循证评价数据库,其中包含近数十万例基于主动报告的中药安全性集中检测数据与中医药临床试验数据,基于此开展了中药安全性和有效性研究,并发布了中药注射剂有效性证据指数^[15]。

4 存在问题

随着信息化与数字化的快速发展,医疗信息数据逐渐从线下转为线上、简单转为复杂、独立转为共享,使医疗领域大数据研究成为必然。由此所带来的不仅仅是医疗数据量的急速膨胀,数据的维度、标准化、结构化等都带来了巨大挑战。虽然当前基于

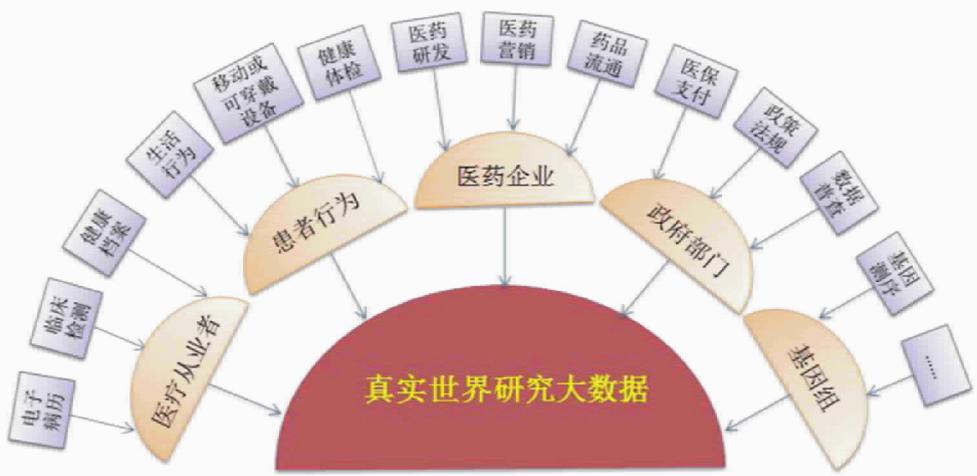


图1 真实世界研究大数据主要来源

大数据的真实世界研究取得了一系列的成果,但是仍然存在着一些问题。

4.1 缺乏多源信息整合平台 “数据孤岛”是指组织中与组织的其他部分隔离且无法访问的信息集合。在当前的真实世界研究中,我国的临床数据极为丰富,但是由于规则和政策机制约束导致数据难以形成共享,出现了“数据孤岛”,乃至存在“不愿共享”“不敢共享”“不能共享”的难题^[16]。由于“数据孤岛”的存在,“各个孤岛”之间的医疗大数据存储结构不一致、存储标准不统一,无疑又增加了数据合并与整合的难度。同时由于不同“数据孤岛”存在地区与政策的差异,极大地影响了真实世界数据的收集质量,高质量数据不能得到保障^[17]。因此打破“数据孤岛”的一个重要前提就是建立真实世界医疗大数据的信息标准,同一信息记录标准的制定有助于数据之间的流动与共享。

4.2 数据规范化程度不高 随着计算机的普及,电子病历系统、临床诊疗系统、医保系统等医疗电子系统广泛得到应用与推广,使得医疗系统大数据完成了初步的数字化存储。然而,不同的数据存储于与不同软件开发的系统之中,字段类型不同、数据编码不一致导致数据规范化程度不高,数据难以联通。还有数据没有完全做到数字,仅仅是图片形式存储。因此数据化程度不高导致大数据可能出现“名不符实”的情况,缺失的部分电子数据可能会对分析结果有重要的影响^[18]。

4.3 复杂数据分析缺乏有效方法 相对于结构化的数据,大数据中充斥了非结构化数据与半结构化数据,结构化数据只占其中很小的一部分。聚焦到真实世界研究中,由于患者个体条件、医疗卫生环境等等因素的影响,导致真实世界研究中的大数据复杂度极大的增加,但是当前在真实世界研究中的大数据技术都是基于已有成熟的框架与算法,例如 Hadoop 和 Map Reduce,缺乏结合医药本身特点进行大数据分析算法的二次开发,没有形成具有医药特色的大数据分析技术,在借用的框架上分析的结果,其信息挖掘的深度与完整性不能得到保障^[19]。

4.4 安全性与隐私性保障问题 在信息化与网络化时代,以互联网为依托的大数据在存储、传输、处理等过程中都会面临着诸多的安全风险^[20]。医疗领域大数据信息多,价值大,并且包含着大量个人隐私数据,因此对安全性与隐私性保障的需求较高,一旦出现数据泄露等安全问题,就会导致巨大的损失或安全威胁^[21-22]。

5 应对策略

中医药的有效性和安全性需要证据支持,这种需求变得越来越旺盛,逐渐成为中医药行业需要突破的瓶颈。基于大数据的真实世界研究为中医药临床评价提供了新的思路和方法,有利于应对中医药快速增长的多样化证据需求的挑战^[23]。针对上述存在的问题,我们提出了相应的建议与应对策略。

5.1 政府部门加快政策制定 基于大数据的真实世界研究为政府部门药品注册审批、医保支付及健康促进等卫生政策制定提供依据。针对当前研究中出现的“数据孤岛”、缺乏多源信息整合平台等问题,政府应当加以重视并加快制定相关政策,包括但不限于数据质量保证、多源数据仓库建设、个人隐私数据保障等等,逐步引导与促进数据的流动与共享,降低重复信息的获取、存储、解析成本,形成大数据驱动下的决策机制。

5.2 加大专项研究投入 政府部门以及相关科研院所应当重视当前中医药大数据及真实世界研究中面临的难题,设立研究专项、加大研究投入、加强学科交叉融合,培育真实世界大数据研究团队,形成基于中医药大数据的真实世界研究体系和规范,推动真实世界研究的不断发展。

5.3 重视人才培养 真实世界大数据即使经过处理之后,其中蕴含的信息与变量也是非常多的,想从浩繁的数据中快速总结规律、挖掘特征、发现问题,需要跨学科的专业人才,需要培养一批具备真实世界研究、大数据分析、数据统计等交叉学科知识的复合型人才。可以通过院校合作、交叉培养等方法,在真实世界研究与大数据研究过程中,加强能力锻炼和知识运用。

真实世界研究概念在引入我国之后得到了快速的发展,为卫生政策制定、临床评价提供了新的思路和方法。随着中医药大数据的不断积累,中医药真实世界研究也在增加,但还处于初级阶段,也面临多种问题。随着政府的大力支持、技术方法的不断进步,基于中医药大数据的真实世界研究也将不断发展完善,成为真实世界证据的重要来源。

参考文献

- [1] 金昌晓,计虹,席韩旭,等. 大数据科研分析平台在临床医学研究中的应用探讨[J]. 中国数字医学,2019,14(2):37-39.
- [2] 杨薇,崔英子,杨海森,等. 医疗大数据在中医药研究领域的应用与思考[J]. 长春中医药大学学报,2016,32(3):625-627.
- [3] Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence-What Is It and What Can It Tell Us? [J]. N Engl J Med,2016,375(23):2293-2297.

- [4] 马金辉, 王志飞, 王永炎, 等. 真实世界大数据 30034 例高血压病住院患者中西医诊疗规律初探[J]. 中国中药杂志, 2014, 39(18):3435-3441.
- [5] 李金根, 姜众会, 高铸辉, 等. 真实世界研究在中医药临床研究中的应用[J]. 世界科学技术-中医药现代化, 2017, 19(1):78-82.
- [6] Makady A., de Boer A., Hillege JL., Klungel O., Goetsch W. What is real-world data? a review of definitions based on literature and stakeholder interviews[J]. Value Health, 2017, 20:858-65.
- [7] Berger ML., Dreyer N., Anderson F., et al. Prospective Observational Studies to Assess Comparative Effectiveness; The ISPOR Good Research Practices Task Force Report[J]. Value Health, 2012, 15(2):217-230.
- [8] 中国大数据产业观察[EB/OL]. (2018-11-15) [2019-10-15]. http://www.cbdio.com/BigData/2018-11/15/content_5918101.htm.
- [9] 中国中医药数据库[EB/OL]. (2019-7-9) [2019-10-15]. <http://cintmed.cintcm.com/cintmed/main.html>.
- [10] 盖国忠. 大数据是中医药真实世界研究的重大机遇——推荐《中医药大数据与真实世界》[J]. 世界中医药, 2019, 14(2):319.
- [11] 刘保延. 真实世界的中医临床科研范式[J]. 中医杂志, 2013, 54(6):451-455.
- [12] 中医大数据中心[EB/OL]. (2018-2-18) [2019-10-15]. <http://www.tcmbigdata.cn/index.html>.
- [13] 中国中药数据中心[EB/OL]. (2015-5-12) [2019-10-15]. <http://www.ndctcm.org/>.
- [14] 人口健康科学数据仓储[EB/OL]. (2019-7-4) [2019-10-15]. [tp://www.ncmi.cn](http://www.ncmi.cn).
- [15] 中药注射剂临床有效性证据指数首次发布[EB/OL]. (2018-10-15) [2019-10-15]. <http://scitech.people.com.cn/n1/2018/1015/c1057-30341108.html>.
- [16] 梁文娜, 林雪娟, 俞洁, 等. 真实世界的大数据助推中医健康管理进入人工智能时代[J]. 中华中医药杂志, 2018, 33(4):1213-1215.
- [17] Real-world evidence: From activity to impact in healthcare decision making[EB/OL]. (2019-9-1) [2019-10-15]. <https://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/real-world-evidence-from-activity-to-impact-in-healthcare-decision-making#>.
- [18] 布优祥, 陈吉, 孙月, 等. 真实世界研究主题与研究能力的可视化分析[J]. 中国中药杂志, 2019, 44(8):1674-1681.
- [19] Benedikt E. Maissenhaelter. Real-world evidence research based on big data[J]. Der Onkologe, 2018, 24:91-98.
- [20] 王晓丽, 万生芳, 魏昭晖, 等. 真实世界大数据下中医临床研究特点探析[J]. 世界最新医学信息文摘, 2019, 19(6):244 + 246.
- [21] 郭杰, 刘善荣. 检验医学真实世界研究的策略与展望[J]. 中华检验医学杂志, 2019, 42(8):618-622.
- [22] 宣建伟, 程江, 薛雄峰, 等. 真实世界医疗大数据的构建及其在医院管理、临床诊疗、合理用药、医保精细化管理中的应用[J]. 中国药物经济学, 2019, 14(5):10-17.
- [23] 杜春霖, 李晓松, 刘元元. 真实世界研究及国内文献综述[J]. 中国卫生信息管理杂志, 2018, 15(5):597-601.
- (2019-11-10 收稿 责任编辑:徐颖)
- (上接第 3118 页)
- [28] 李和伟, 王启帆, 付宇, 等. 中医防控慢性病的优势与发展思考[J]. 中华中医药杂志, 2018, 33(11):4821-4823.
- [29] 郭彧, 孙李李, 谭云龙, 等. 中国慢性病前瞻性研究的标准生物银行建设[J]. 转化医学杂志, 2014, 3(6):321-326.
- [30] 胡智超, 葛利达, 柳涛, 等. 真实世界研究中如何关注中西药相互作用的安全性评价[J]. 中国循证医学杂志, 2018, 18(11):1158-1162.
- [31] 谢雁鸣, 王连心. 中西成药联合应用, 临床疗效精准定位[J]. 中国中药杂志, 2016, 41(24):4479-4482.
- [32] Skalli S, Bencheikh R S. Safety monitoring of herb-drug interactions[J]. Drug safety, 2012, 35(10):785-791.
- [33] Vohra S, Cvijovic K, Boon H, et al. Study of natural health product adverse reactions (SONAR): active surveillance of adverse events following concurrent natural health product and prescription drug use in community pharmacies[J]. PloS one, 2012, 7(9):e45196.
- [34] 杨会生, 谢雁鸣, 陈岑, 等. 基于关联规则分析的“真实世界”复方苦参注射液治疗肺癌中西药联合应用特征研究[J]. 中国中药杂志, 2018, 43(8):1708-1713.
- [35] 王处渊, 谢雁鸣, 王连心, 等. 真实世界 25338 例肠恶性肿瘤患者中西药联合应用特征分析[J]. 中国中医基础医学杂志, 2015, 21(12):1544-1549.
- [36] 张文涛, 韩永鹏. 中西结合药学的研究进展[J]. 临床医药文献(连续型电子期刊), 2018, 5(A0):216-218.
- [37] 郑文科, 张莉, 管树杰. 中药饮片亟需开展临床评价[J]. 中国中药杂志, 2019, 44(3):624-628.
- [38] 肖永庆, 李丽, 刘颖. 构建中药饮片质量保障体系的关键问题[J]. 世界科学技术-中医药现代化, 2015, 17(1):167-172.
- [39] 冯国栋. 浅谈可影响中药饮片临床疗效的相关因素[J]. 当代医药论丛, 2018, 16(11):44-45.
- [40] 黄卓山, 罗艳婷, 刘金来. 真实世界研究的方法与实践[J]. 循证医学, 2014, 14(6):364-368.
- [41] 田磊, 岳彩宾, 管欣, 等. 真实世界研究与随机对照试验在临床实践及卫生决策中应用的比较[J]. 中国医院药学杂志, 2019, 39(3):274-277.
- [42] 邹冲, 蒋萌, 蒋卫民, 等. 构建基于临床大数据真实世界的中医药临床科研模式[J]. 中华中医药杂志, 2016, 31(10):4113-4116.
- [43] Zhang J, Zhang B. Clinical research of traditional Chinese medicine in big data era[J]. Frontiers of medicine, 2014, 8(3):321-327.
- (2019-11-10 收稿 责任编辑:徐颖)