·新药述评与论坛·

真实世界数据存在的问题与质量提升对策研究

李宛亭[1,2],乔佳慧[1,2],孟令全[1,2],连桂玉[1,2],黄　哲[1,2],陈玉文[1,2]
(沈阳药科大学 1 工商管理学院;2 药品监管科学研究院, 沈阳 110016)

[摘要] 传统的随机对照临床试验在长期使用过程中暴露出一些不足,这促使越来越多的研究者将目光投向真实世界证据(RWE)。真实世界数据(RWD)作为真实世界研究(RWS)的重要基础,决定了 RWS 是否能为监管和审评机构提供可靠的 RWE。目前 RWD 在使用过程中存在一些问题,表现为数据非结构化和零碎化、可及性差、数据的标准化处理缺乏数据标准、数据收集所依据的研究设计不严谨、未严格把控 RWD 的适用性等,这些问题极大地影响 RWD 的质量。本文针对以上问题,提出加强对医疗机构中的 RWD 的管理、完善 RWD 的共享机制、RWS 过程中应根据数据标准将 RWD 标准化、RWD 的收集应依据严谨的研究设计、明确与 RWD 质量相关人员的职责、严格把控 RWD 的适用性 6 条建议,以期解决 RWS 中 RWD 存在的质量问题,为提升 RWD 质量提供参考。

[关键词] 真实世界研究;真实世界证据;真实世界数据;数据质量
[中图分类号] R95　[文献标志码] A　[文章编号] 1003 - 3734(2021)13 - 1160 - 04

Research on the problems of real-world data and the measures to improve real-world data's quality

LI Wan-ting[1,2], QIAO Jia-hui[1,2], MENG Ling-quan[1,2], LIAN Gui-yu[1,2], HUANG Zhe[1,2], CHEN Yu-wen[1,2]
(1 *College of Business Administration*, 2 *Research Institute of Drug Regulatory Science*, *Shenyang Pharmaceutical University*, *Shenyang* 110016, *China*)

[**Abstract**] Traditional clinical trials that are randomized and controlled have exposed some deficiencies in the process of long-term use, prompting a growing number of researchers to look at real-world evidence (RWE). As an important basis of real-world study (RWS), real-world data (RWD) determine whether the RWS is qualified for supervision and review of institutions with reliable RWE. There are some problems about RWD, including unstructured and fragmented data, poor accessibility, lack of standards to standardize data, research design that is not rigorous, failure to control the applicability of RWD, etc. These problems will greatly affect the quality of RWD. Aiming to resolve the above problems, this paper provides six suggestions in order to provide a reference to improve the quality of RWD, including strengthening the management of RWD in medical institutions, improving RWD sharing mechanism, standardizing RWD according to data standards in RWS, collecting RWD according to rigorous research design, defining the responsibilities of personnel related to RWD quality and strictly controlling the applicability of RWD.

[**Key words**] real-world study; real-world evidence; real-world data; quality of data

在制定临床指南和药品审评决策中,传统的随机对照临床试验(random controlled trial, RCT)一直被视作是“金标准”,其结果也被许多循证医学家用来作为较高标准等级的临床证据。但其也存在一些

[基金项目] 国家药品监督管理局药品监管科学研究基地-沈阳药科大学药品监管科学研究院专项基金(2020jgkx005)
[作者简介] 李宛亭,女,硕士研究生,主要从事医药制造业 R&D 效率研究、研发投入与风险研究。E-mail:1053923048@qq.com。
[通讯作者] 陈玉文,男,教授,博士生导师,主要从事医药制造业 R&D 效率研究、研发投入与风险研究。联系电话:(024)23986552,E-mail:cywwyc@163.com。黄哲,女,教授,博士生导师,主要从事药事管理学研究。E-mail:huangzhe2000@sina.com。

问题,如纳排标准比较严苛,结果的外推性较差,研究成本较高,或是在一些特殊情况下(如罕见病治疗),出于伦理考虑或因招募困难等无法开展研究。基于这样的背景,许多国家的监管机构开始尝试用真实世界研究(real-world study,RWS)来弥补 RCT 的不足,并发布了与 RWS 相关的指导文件。我国国家药品监督管理局也分别于 2019 年 12 月与 2020 年 1 月发布了《真实世界数据用于医疗器械临床评价技术指导原则(征求意见稿)》[1]和《真实世界证据支持药物研发与审评的指导原则(试行)》[2],从而正式拉开了将 RWS 应用于支持药械研发、审评与监管的序幕。除了政策上的支持,大数据平台与信息技术的不断发展也为 RWS 提供了机遇。真实世界数据(real-world data, RWD)是进行 RWS 的关键,其质量的高低决定其能否成为真实世界证据(real-world evidence, RWE),并直接影响着证据的强度。因此,如何提高 RWD 的质量成为增强 RWE 可信度以及推动 RWS 广泛应用的关键。

1 相关概念

《真实世界证据支持药物研发与审评的指导原则(试行)》[2]指出,RWS 是指基于预设的临床问题,对在真实世界环境下收集到的与研究问题相关的 RWD 进行合理分析,从而得到 RWE 的研究过程;RWD 是指来源于日常所收集的与患者健康状况和/或诊疗及保健有关的数据;RWE 是指通过对 RWD 进行适用性评估和合理的分析而得到的药物使用情况以及获益-风险的临床证据。

临床试验指在人体进行药物的系统性研究,目的是确定试验药物的疗效与安全性[3]。临床试验过程中产生的数据被称为临床试验数据。

2 RWD 的来源

按照数据的获取方式来分,RWD 可来源于回顾性数据库里的数据和依照研究目的前瞻性收集到的数据[4]。回顾性数据库是指早就存在的、并非针对特定研究目的而建立的数据库[5]。回顾性数据库有时也被表述为既有健康医疗数据,其中包含大量的数据,包括电子健康档案数据、疾病登记数据、医保索赔数据、出生/死亡登记数据等[2]。仅依靠回顾性数据库中的数据难以支持 RWS,许多时候存在数据不准确、不完整等问题,这种情况下就需要针对研究目的进行预先的设计,在多种质量控制措施下主动收集相关数据,即前瞻性收集的数据,如患者登记数据、实用临床试验数据等。

3 RWD 存在的问题

RWE 与传统临床研究证据最本质的区别在于数据获取的环境不同[6]。RWS 的入组人群标准比较宽泛,能收集到真实临床、社区、家庭等多样化的 RWD,相较于传统的随机对照临床试验而言更贴近实际[7]。但目前 RWD 仍存在诸多问题:① 由于医疗机构中的日常就诊数据、临床试验数据等分布广而无序,所能提供的信息也较为碎片化,加之缺乏统一的数据标准,多为非结构或者半结构数据,导致挖掘这些数据所得到的原始数据质量较差[8],研究人员在基于这些数据做回顾性研究时往往需要对数据进行繁琐的清洗和复杂的处理,增加了工作负担。② 基于数据保护的目的,目前国内的医疗数据开放程度还比较小,与英、美等国家相比还存在较大的差距[9],不利于 RWD 的可及性和完整性。③ 在进行真实世界研究设计时,由于研究者对一些变量的定义或者结局指标拿捏不准,导致设计缺乏严谨性,依据研究设计收集所得的 RWD 准确性欠佳。④ 由于真实世界研究一般是在更大范围的人群中进行,可能会存在混杂因素或者偏倚问题等。⑤ 申办人与研究者在真实世界研究过程中未把控好 RWD 的适用性评估,导致在转化为高质量的 RWE 时缺乏合理的依据。RWD 所存在的这些问题极大地影响着数据质量,也限制了 RWS 的应用,因此提高 RWD 质量显得尤为重要。

4 RWD 质量提升对策

4.1 加强对医疗机构中的数据管理 医疗机构是产生健康医疗数据以及临床试验数据最重要的场所,如若能够从源头上加强医疗机构中的数据的管理,则能够大大提高所提取或收集 RWD 的质量以及研究效率。首先要推动医疗机构与医生深化数据电子化的理念,相对于电子化数据,纸质数据的收集较为滞后、资料整理与保存麻烦且难以达到同步更新和核查的要求。因此应尽量使用电子病历或电子病例报告表,方便数据的收集、核查与修正[10]。对于临床试验数据,尤其是以注册为目的的临床试验,应遵循一定的质量标准,多个国家的监管机构在发布的指导原则中提到将 ALCOA + CCEA 标准作为临床试验数据质量的国际通行原则[11]。此外,临床试验数据的管理还应严格遵循国家药监部门发布的《临床试验数据管理工作技术指南》[12]、《临床试验的电子数据采集(EDC)技术指导原则》[13]、《药物临床试验质量管理规范》[14]等相关文件的要求。对

于既有的健康医疗数据，国家相关部门应出台政策规范，统一规范医疗机构对健康医疗数据的分类、核查、清理、存储、访问权限等，使可能影响数据质量的各种因素得以控制，并便于研究者对既有健康医疗数据的采集与利用。

4.2 完善 RWD 的共享机制 许多医疗机构、政府部门基于数据安全的考虑，对一些医疗、医保数据等开放程度小。数据资源的开放共享对于研究项目至关重要，国家相关部门应对来自不同区域或不同部门的数据进行分类与存储，并在去除患者的个人敏感信息的基础上，建立数据资源共享交换平台；为了实现不同医疗机构、不同政府部门、不同区域以及不同领域之间数据资源共享互通，还应依托现有技术加快完善分类别、分级别、分部门、分地域的健康医疗数据共享机制，从而消除信息孤岛问题，让研究者在确保数据安全的条件下有规范可依地获取完整和准确的相关数据。

4.3 RWS 过程中应根据数据标准将 RWD 标准化 目前医疗机构信息系统中存在的健康医疗数据以及临床试验数据大多是非结构化和零散的，为了能够更快地将不同科室和不同业务系统的非结构化数据整合为高质量数据，以及实现这些数据与临床研究数据采集系统之间的数据互通共享，并且便于申办人与审评和监管机构之间的交流，应对这些既有的健康医疗数据或者临床试验数据进行标准化。如数据或信息的交换标准可参照 CDISC 标准、HL7 标准等，医学术语及代码标准可参照 MedDRA、世界卫生组织药物词典、WHOART 术语集、ICD-10 国际疾病分类标准编码等，已有的医疗数据和临床试验数据可按照上述标准进行标准化设计与实施。

4.4 RWD 的收集应依据严谨的研究设计 RWD 的收集是围绕研究设计而进行的，药品审评机构以及监管机构在进行审评和决策时也极为看重研究设计所得的证据能否回答和解决相关的临床问题[15]。在 RWS 中，常见的研究设计类型有观察性研究、实用临床试验、以 RWD 作为外部对照的单臂临床试验。为了把控数据结果的质量，研究者在进行研究设计时应当重点把控数据收集的可行性，明确研究人群的纳排标准，保证所涉及的重要数据要素和影响因素变量的定义、识别或算法准确，以及合理建立数据的清理和核查准则。不论哪种研究设计，研究者还应选择合适的统计学方法对收集到的数据中的偏倚、混杂因素等风险进行控制，对缺失值、异常值等进行处理，如倾向评分法、疾病风险评分法、工具变量法、敏感性分析等。

4.5 明确与 RWD 质量相关人员的职责 数据的质量控制离不开相关人员的参与。与 RWD 质量控制相关的人员包括申办者、研究者、监查员、临床协调员、数据管理员以及合同研究组织等[12]。申办者对 RWD 质量负总责，应对 RWD 质量进行管理、检查和监督。研究者应及时提交数据，并确保数据的准确性和完整性。监查员应依照原始文档对 RWD 进行核查，确保数据的记录完整、正确。涉及多中心的临床试验的，还需要安排临床协调员与各个中心的医生充分沟通，并将医生对数据的评价、建议与反馈汇报给研究小组。数据管理员应熟悉研究设计，建立为研究所服务的研究型数据库，对数据标准和逻辑进行管理与检验。当涉及收集临床试验数据时，通常会与合同研究组织合作，合同研究组织在收集数据的过程中应该严格按照操作流程，并对数据质量进行控制。在 RWS 过程中，应对这些与 RWD 质量相关的人员进行规范的培训与监督，使其明确自身职责。

4.6 严格把控 RWD 的适用性 适用性评估是检验 RWD 是否能够转化为高质量的 RWE 的标准。满足适用性的 RWD 才可能为审评和监管机构提供可信度高的 RWE。因此从数据的采集到管理都应严格评估 RWD 的适用性。适用性有 2 个维度，包括相关性和可靠性[2]。相关性是指 RWD 应与临床问题相关，包含与临床结局相关的变量，有明确的临床结局定义和临床意义，研究人群应具代表性，并且有足够的样本量和随访时间，这反映出审评和监管机构对 RWE 所能解决的临床问题的重视和关注。可靠性是指 RWD 在与临床问题相关的基础上，还应保证准确、完整和可溯源，同时，在数据的收集与管理过程中应有质量保证措施。

5 结语

随着我国有关 RWS 的指导原则出台，RWD 与 RWE 将对医疗卫生领域产生巨大的影响。国家相关部门应加大对出台的相关文件的落实效果进行监管，规范医疗机构对健康医疗数据的管理，完善 RWD 共享机制，加快 RWD 共享开放平台的建立，并为数据的标准化提供可靠的数据标准。研究者和申办者在进行 RWS 时，要遵循国家相关指导文件的要求，充分利用现有的数据资源以及先进技术，科学严谨地进行研究设计并依据设计收集 RWD，明确与

RWD 质量相关人员的职责,全程把控好 RWD 的适用性,以期对影响数据质量的因素加以控制,提高数据质量,为监管和审评机构做出决策提供可靠和高质量的 RWE,从而发挥出 RWS 真正的优势与作用。

[参 考 文 献]

[1] 国家药品监督管理局. 真实世界数据用于医疗器械临床评价技术指导原则(征求意见稿)[EB/OL]. (2019 - 12 - 13). https://www.cmde.org.cn/CL0066/20139.html.

[2] 国家药品监督管理局. 真实世界证据支持药物研发与审评的指导原则(试行)[EB/OL]. (2020 - 01 - 03). http://www.nmpa.gov.cn/WS04/CL2138/373175.html.

[3] 临床试验[EB/OL]. [2012 - 12 - 06]. https://baike.baidu.com/item/%E4%B8%B4%E5%BA%8A%E8%AF%95%E9%AA%8C/7626892?fr=aladdin.

[4] 王雯,刘艳梅,谭婧,等. 回顾性数据库研究的概念、策划与研究数据库构建[J]. 中国循证医学杂志,2018,18(2):230 - 237.

[5] MOTHERAL B, BROOKS J, CLARK MA, *et al*. A checklist for retrospective database studies—report of the ISPOR task force on retrospective databases[J]. *Value Health*, 2003, 6(2): 90 - 97.

[6] 刘晓清,孙晓川. 真实世界证据[J]. 协和医学杂志,2017,8(Z2):305 - 310.

[7] 戴亮,季光. 基于 FDA《RWE 计划框架》论析 RWS 对中药新药研发的意义[J]. 中药新药与临床药理,2019,30(11):1403 - 1408.

[8] 常朝娣,陈敏. 大数据时代医疗健康数据治理方法研究[J]. 中国数字医学,2016,11(9):2 - 5.

[9] 赵飞,武琼,曾良怀,等. 健康医疗大数据共享开放模式研究[J]. 中国卫生信息管理杂志,2019,16(6):649 - 654.

[10] 温泽淮,李玲,刘艳梅,等. 实效性随机对照试验的技术规范[J]. 中国循证医学杂志,2019,19(7):794 - 802.

[11] 董冲亚,姚晨,高嵩,等. 加强医院临床研究源数据管理,提高我国临床研究数据质量[J]. 中国循证医学杂志,2019,19(11):1255 - 1261.

[12] 国家食品药品监督管理局. 临床试验数据管理工作技术指南[EB/OL]. (2016 - 07 - 29). http://www.nmpa.gov.cn/WS04/CL2138/300192.html.

[13] 国家食品药品监督管理局. 临床试验的电子数据采集(EDC)技术指导原则[EB/OL]. (2016 - 07 - 29). http://www.nmpa.gov.cn/WS04/CL2182/300194.html.

[14] 国家药品监督管理局. 药物临床试验质量管理规范[EB/OL]. (2020 - 04 - 23). http://www.nmpa.gov.cn/WS04/CL2138/376852.html.

[15] 姚晨. 利用好 RWD 生产高质量 RWE 支持药械监管[J]. 中国食品药品监管,2020(2):22 - 27.

编辑:王宇梅/接受日期:2020 - 10 - 09