

真实世界证据与随机对照试验: RCT DUPLICATE 项目启动、实施、进展解读 与启示(二)

石舒原^{1,2} 刘佐相¹ 赵厚宇¹ 聂晓璐^{1,3} 韩晟⁴ 符祝⁵ 宋海波^{6,7} 姚晨^{5,8}
詹思延^{1,9} 孙凤^{1,5}

¹北京大学公共卫生学院流行病学与卫生统计学系,北京 100191;²中国康复科学所,中国残联残疾预防与控制中心,北京 100068;³国家儿童医学中心,首都医科大学附属北京儿童医院临床流行病学与循证医学中心,北京 100045;⁴北京大学药学院药事管理与临床药学系,北京 100191;⁵海南省真实世界数据研究院,博鳌乐城国际医疗旅游先行区管理局,乐城 571437;⁶国家药品监督管理局药品评价中心,北京 100022;⁷国家药品监督管理局药物警戒研究与评价重点实验室,北京 100022;⁸北京大学临床研究所,北京 100191;⁹北京大学第三医院临床流行病学研究中心,北京 100191

通信作者:孙凤,Email: sunfeng@bjmu.edu.cn

【摘要】 随着医疗大数据的推广与应用,监管机构开始使用非随机化的真实世界证据(RWE)来评估医疗产品的有效性。本文简要介绍了哈佛大学研究团队于2018年发起RCT DUPLICATE项目的最新进展与研究成果,并结合我国医疗服务特点总结了相关研究经验,为国内学者今后开展相关RWE研究提供借鉴。

【关键词】 随机对照试验; 真实世界证据; 真实世界数据

基金项目: 国家自然科学基金(72074011); 中国药品监管科学行动计划第二批重点项目([2021]37-10); 海南省博鳌乐城先行区真实世界研究专项计划(HNLG2022RWS012)

Real-world evidence and randomized controlled trials: the initiation, implementation, progress interpretation and revelation of RCT DUPLICATE (part 2)

Shi Shuyuan^{1,2}, Liu Zuoxiang¹, Zhao Houyu¹, Nie Xiaolu^{1,3}, Han Sheng⁴, Fu Zhu⁵, Song Haibo^{6,7}, Yao Chen^{5,8}, Zhan Siyan^{1,9}, Sun Feng^{1,5}

¹Department of Epidemiology and Biostatistics, Peking University School of Public Health, Beijing 100191, China; ²China Institute of Rehabilitation Sciences, Center for Prevention and Control of Disability of China Disabled Persons Federation, Beijing 100068, China; ³Center for Clinical Epidemiology and Evidence-based Medicine, Beijing Children's Hospital, Capital Medical University, National Center for Children's Health, Beijing 100045, China; ⁴Department of Pharmacy Management and Clinical Pharmacy, Peking University School of Pharmacy, Beijing 100191, China; ⁵Hainan Institute of Real World Data, the Administration of Boao Lecheng International Medical Tourism Pilot Zone, Lecheng 571437, China; ⁶Center for Drug Reevaluation, National Medical Products Administration, Beijing 100022, China; ⁷Key Laboratory for Research and Evaluation of Pharmacovigilance, National

DOI: 10.3760/cma.j.cn112338-20220513-00409

收稿日期 2022-05-13 本文编辑 万玉立

引用格式: 石舒原, 刘佐相, 赵厚宇, 等. 真实世界证据与随机对照试验: RCT DUPLICATE 项目启动、实施、进展解读与启示(二)[J]. 中华流行病学杂志, 2022, 43(11): 1835-1841. DOI: 10.3760/cma.j.cn112338-20220513-00409.

Shi SY, Liu ZX, Zhao HY, et al. Real-world evidence and randomized controlled trials: the initiation, implementation, progress interpretation and revelation of RCT DUPLICATE (part 2)[J]. Chin J Epidemiol, 2022, 43(11): 1835-1841. DOI: 10.3760/cma.j.cn112338-20220513-00409.



Medical Products Administration, Beijing 100022, China; ⁸ Peking University Clinical Research Institute, Beijing 100191, China; ⁹ Clinical Epidemiology Research Center, Peking University Third Hospital, Beijing 100191, China

Corresponding author: Sun Feng, Email: sunfeng@bjmu.edu.cn

【 Abstract 】 With the promotion and application of big medical data, non-interventional real-world evidence (RWE) has been used by regulators to assess the effectiveness of medical products. This paper briefly introduces the latest progress and research results of the RCT DUPLICATE Initiative launched by the research team of Harvard University in 2018 and summarizes relevant research experience based on the characteristics of China's medical service to provide inspiration and reference for domestic scholars to conduct related RWE research in the future.

【 Key words 】 Randomized controlled trial; Real-world evidence; Real-world data

Fund programs: National Natural Science Foundation of China (72074011); The Second Batch of Key Projects of Scientific Act for Drug Regulation of China ([2021]37-10); Real World Study Project of Hainan Boao Lecheng Pilot Zone (HNLC2022RWS012)

药物治疗的有效性和安全性相关的医疗卫生监管决策需要真实的证据作为依据,随机对照试验(randomized controlled trial, RCT)虽被视为证据来源的金标准^[1],但临床决策不能仅依靠 RCT 的证据,仍需要真实世界证据(real-world evidence, RWE)的补充支持。近些年,利用真实世界数据(real-world data, RWD)产生的 RWE 被越来越多地用于补充 RCT 的决策范畴,包括监测疾病状况、评估药物疗效和支持监管决策^[2-5]等。但目前关于 RWE 研究的探索和应用仍面临着众多挑战,包括 RCT 与 RWE 研究结果是否一致, RWE 研究何时可以替代 RCT, 如何更好地利用 RWE 用于医疗领域等。

2016 年出台的《21 世纪治愈法案》要求美国食品药品监督管理局(FDA)促进 RWE 的使用^[6]。2018 年 FDA 正式发布了 RWE 项目的实践框架^[7],列出了 FDA 后续行动计划的实施细节,包括示范项目、涉及的利益相关者以及特定主题的指导文件,以帮助发展 RWE,支持 FDA 的监管决策。欧洲药品管理局在对开展 RWE 研究实践活动保持警惕的同时,也在持续推动 RWE 在决策方面的应用^[8]。2019 年 5 月国家药品监督管理局药品评价中心也发布了使用 RWE 支持药物开发关键考虑的指导意见草案^[9]。

RCT DUPLICATE (Randomized, Controlled Trials Duplicated Using Prospective Longitudinal Insurance Claims: Applying Techniques of Epidemiology)项目便是鼓励在监管决策层面使用 RWE 的举措之一,旨在证明 RWE 在何种情况下可以模拟 RCT 的结果。该项目将为 RWE 的因果推论提供参考案例,以增加使用 RWE 进行监管决策的

可信度。基于此,本课题组撰写了以 RCT DUPLICATE 项目为基础的系列文章^[10-12]。并持续跟进该项目的进展,在系列一中对 RCT DUPLICATE 项目背景和研究内容框架进行概述的基础上,本文主要介绍项目进展与主要成果,并对两个具有代表性研究的实施、进展及成果进行重点解读。本文总结了该项目对我国开展相关 RWE 研究带来的一些启示,为国内学者今后开展 RWE 研究提供借鉴。

一、RCT DUPLICATE 项目研究进展

1. 目前已公开的研究结果:本文针对项目组最新公布的 10 个 RWE 研究结果展开介绍^[13],基本情况见表 1。其中 8 项为降糖药比较,2 项为抗血小板药比较,考虑的临床结局均为心血管相关结局。其中 7 项 RCT 为安慰剂对照,项目组在模拟 RWE 中选择的阳性对照组来替代安慰剂组,因为未经治疗的患者可能从本质上区别于主动治疗的患者,而这些个体差异很难在医疗索赔数据中捕捉到。项目组规定,在 GLP-1 受体激动药(GLP-1RA)和 SGLT-2 抑制剂(SGLT-2i)的研究中,使用 DPP-4 抑制剂(DPP-4i)作为对照组,在 DPP-4i 的研究中,使用磺脲类药物作为对照组。之所以选择 DPP-4i 和磺脲类药物作为安慰剂的替代药物,因为它们既是抗糖尿病药物,又与研究的药物有相似的适应症,且有证据表明,它们对研究关注的心血管结局没有影响。

在 10 个研究中, RWE 研究和 RCT 的结果差异无统计学意义。根据事先制定的评价标准^[26],其中 8 个 RWE 研究的结果与对应的 RCT 结果实现了估计值一致性,6 个 RWE 研究的结果与对应的 RCT 结果实现了监管一致性(RWE 研究模拟 RCT 结果的方向和统计显著性的能力)。这些研究结果也指

表 1 10 项 RWE 研究的基本介绍

简称	NCT 号码	类型	干预药物(对照)		结局
			随机对照试验	真实世界证据	
LEADER ^[14]	NCT03936049	糖尿病药	利拉鲁肽(安慰剂)	利拉鲁肽(DPP-4i)	3P MACE
DECLARE ^[15]	NCT04215523	糖尿病药	达格列净(安慰剂)	达格列净(DPP-4i)	心衰入院+心血管死亡
EMPA-REG ^[16]	NCT04215536	糖尿病药	恩格列净(安慰剂)	恩格列净(DPP-4i)	3P MACE
CANVAS ^[17-18]	NCT03936010	糖尿病药	卡格列净(安慰剂)	卡格列净(DPP-4i)	3P MACE
CARMELINA ^[19]	NCT03936036	糖尿病药	利格列汀(安慰剂)	利格列汀(磺脲类药物)	3P MACE
TECOS ^[20]	NCT03936062	糖尿病药	西他列汀(安慰剂)	西他列汀(磺脲类药物)	3P MACE+心绞痛
SAVOR-TIMI ^[21]	NCT03936023	糖尿病药	沙格列汀(安慰剂)	沙格列汀(磺脲类药物)	3P MACE
CAROLINA ^[22-23]	NCT03648424	糖尿病药	利格列汀(安慰剂)	利格列汀(格列美脲)	3P MACE
PLATO ^[24]	NCT04237935	抗血小板药	替格瑞洛(氯吡格雷)	替格瑞洛(氯吡格雷)	3P MACE
TRITON-TIMI ^[25]	NCT04237922	抗血小板药	普拉格雷(氯吡格雷)	普拉格雷(氯吡格雷)	3P MACE

注:数据来自 <https://www.clinicaltrials.gov>。(更新时间为 2022 年 4 月 12 日);LEADER:利拉鲁肽在糖尿病中的作用:心血管预后评价;DECLARE:达格列净对心血管事件发生率影响的多中心研究;EMPA-REG:2 型糖尿病患者的心血管结局事件试验;CANVAS:卡格列净心血管评价研究;CARMELINA:利格列汀治疗 2 型糖尿病患者的心血管和肾微血管预后研究;TECOS:西他列汀心血管预后研究;SAVOR-TIMI:沙格列汀单独使用或与其他糖尿病药物联合使用是否能降低心血管事件的风险;CAROLINA:利格列汀与格列美脲对 2 型糖尿病患者心血管事件的研究;PLATO:血小板抑制与患者预后;TRITON-TIMI:普拉格雷溶栓优化血小板抑制改善心肌梗死治疗效果的试验;3P MACE:主要不良心血管事件复合结局,包括非致死性心肌梗死、非致死性中风或心血管死亡

出了 RWE 研究的一个主要局限性:由于真实世界中存在诊疗的差异,以及在真实临床环境中接受药物治疗的患者与参加 RCT 的患者之间的本质差异,尤其针对以安慰剂为对照的 RCT 的模拟可能具有挑战性。在未达成监管一致性的 4 个 RWE 研究中,有 3 个对应的 RCT 采用的是安慰剂对照^[13]。每个 RWE 研究的详细信息可在注册网站(<https://www.clinicaltrials.gov>)上查询获取。

(1) 基线特征:10 个 RCT 中研究对象的平均或中位年龄从 61 岁到 66 岁不等,除了 DPP-4i 的 RCT (CARMELINA^[19]、TECOS^[20]、SAVOR-TIMI^[21] 和 CAROLINA^[22]) 的 4 项 RWE 研究(年龄比 RCT 人群更大,均 $P < 0.05$) 外,每个 RWE 研究患者的平均年龄与相应的 RCT 的平均年龄相近。这表明尽管设计 RWE 研究时采取与 RCT 人群相同的年龄范围标准,但 RWE 研究的人群年龄普遍会略大于 RCT 人群的年龄。此外,除 TRITON-TIMI^[25] 外,所有 RWE 研究中女性的占比均高于 RCT 中女性占比。在 RCT 和相应的 RWE 研究中,测量的心血管结局的危险因素是相近的,例如都包括吸烟和高血压等因素。

(2) 结局事件发病密度:10 个 RCT 中除了 DECLARE^[15], 9 个均使用主要不良心血管事件复合结局(3P MACE, 包括非致死性心肌梗死、非致死性中风或心血管死亡) 作为主要终点,其 RWE 研究中的发病密度均较低。例如,在 LEADER 试验的试验

组和对照组中,发病密度分别为 3.4/100 人年和 3.9/100 人年,而对应的 RWE 研究的发病密度在两组分别为 2.0/100 人年和 2.8/100 人年^[14]。然而,有其他结局指标的两个 RCT 在其对应的 RWE 研究中表现出更高的发病密度。如在 DECLARE 试验的试验组和对照组中,因心力衰竭住院或心血管疾病死亡的发病密度分别为 1.2/100 人年和 1.5/100 人年,而其 RWE 研究的发病密度分别为 1.6/100 人年和 2.4/100 人年^[15]。TECOS 试验的 3P MACE 或因急性心绞痛住院的发病密度在试验组和对照组中分别为 4.1/100 人年和 4.2/100 人年,而其 RWE 研究的发病密度分别为 7.3/100 人年和 8.3/100 人年^[20]。这些发病密度的差异可能归因于研究人群本身的差异,但考虑到 RWE 研究中事件捕获的特异性较低,项目组将这些试验的终点评估的质量标记为中等。在随访的前 6 个月,RWE 研究中的患者数量迅速下降,导致 RWE 研究的平均随访时间较短。RWE 研究中的事件数一般是通过大量患者积累得来的,而 RCT 的患者较少,但其随访时间较长。

(3) RCT-RWE 结果一致性:10 个 RWE 研究中^[13], 有 6 个符合监管一致性(LEADER、DECLARE、EMPA-REG、CANVAS、CAROLINA 和 TRITON-TIMI), 有 8 个符合估计值一致性(LEADER、EMPA-REG、CANVAS、CARMELINA、TECOS、CAROLINA、TRITON-TIMI 和 PLATO)。PLATO 结果发现替格瑞洛优于氯吡格雷($HR=0.84, 95\%CI: 0.77-0.92$)^[24],

其 RWE 研究发现 HR 点估计与 RCT 的估计值处在同一方向,但其 95%CI 的上限大于 1.00 ($HR=0.92$, $95\%CI:0.83\sim 1.02$),故该模拟研究满足预先制定的估计值一致性,而未满足监管一致性。比较 DPP-4i 与安慰剂的 3 项 RCT (CARMELINA^[19]、TECOS^[20] 和 SAVOR-TIMI^[21]) 结果发现,在心血管风险方面, DPP-4i 相比于安慰剂是非劣效的而不是优效的,对应的 RWE 研究均呈现出了优效性。有 8 个 RCT 的模拟研究都达到了估计值一致性(除了 DECLARE^[15] 和 SAVOR-TIMI^[21]),即这些 RWE 研究的点估计值都落在了 RCT 估计值的 95%CI 内。此外,项目组目前结果显示 SAVOR-TIMI 是唯一一个 RCT 和 RWE 研究估计值有显著性差异的试验。

2. 具体案例解读:

(1) 模拟已完成的 RCT 项目:模拟 CANVAS 的研究:CANVAS 试验 (NCT01032629, NCT01989754) 包含两项多中心、双盲、安慰剂对照试验 (CANVAS 和 CANVAS-R)^[17],在 30 个国家的 667 个中心开展,旨在评估 2 型糖尿病 (T2DM) 患者中服用卡格列净对发生 3P MACE 的影响。CANVAS 试验研究对象为年龄在 30 岁以上血糖控制不佳 [$7.0\% \leq$ 糖化血红蛋白 (HbA1c) $\leq 10.5\%$] 且有动脉粥样硬化性血管病症状史的患者,或年龄在 50 岁以上且至少有两种心血管危险因素的患者。

试验结果显示研究对象的平均年龄为 63.3 岁, 35.8% 为女性, T2DM 平均病程为 13.5 年, 65.6% 有心血管病史。卡格列净组的主要结局发生率低于安慰剂组 (26.9/1 000 人年 vs. 31.5/1 000 人年, $HR=0.86$, $95\%CI:0.75\sim 0.97$, 非劣效 $P<0.001$, 优效 $P=0.020$)。

模拟 CANVAS 试验的 RWE 队列研究 (NCT03936010) 研究是利用美国医疗保险数据库 (OpumInsight, Eden, Prairie 和 MN) 来评价与临床常规治疗 T2DM 的 DPP-4i、GLP-1RA 和磺脲类药物相比^[18],卡格列净治疗 T2DM 的心血管安全性。研究采用倾向性评分法将卡格列净的新用药者分别与 3 种对照药物的新用药者进行了 1:1 匹配,构建了 3 个研究队列 (卡格列净与 DPP-4i、GLP-1RA 或磺脲类药物匹配后 3 对样本量分别为 17 667 对、20 539 对和 17 354 对)。

研究发现与常规抗糖尿病药物 (DPP-4i、GLP1-RA 和磺脲类) 初次使用者相比,卡格列净初次使用者因心力衰竭入院风险明显降低 ($HR=0.70$, $95\%CI:0.54\sim 0.92$; $HR=0.61$, $95\%CI:0.47\sim 0.78$; $HR=$

0.51 , $95\%CI:0.38\sim 0.67$), 在心肌梗死或中风的复合结局发生方面无明显差异。

(2) 预测未完成的 RCT 项目:模拟 CAROLINA 的研究:CAROLINA 试验 (NCT01243424) 是一项多中心、随机、双盲、阳性对照试验,共纳入 6 042 例患者^[22]。研究起始时间为 2010 年 11 月 11 日,完成时间为 2018 年 8 月 21 日。该研究旨在评估利格列汀与格列美脲对早期 T2DM 患者发生 3P MACE 的影响。此外, CAROLINA 还包含了 4 项子研究:认知子研究 (cognition sub-study)、血糖变异性子研究 (glycaemic variability sub-study)、 β 细胞功能子研究 (β -cell function sub-study) 和成人隐匿性自身免疫性糖尿病子研究 [latent autoimmune diabetes in adults (LADA) sub-study]。

CAROLINA 试验结果显示 6 042 名受试者中, 6 033 名参与了治疗并进入到分析中 (平均年龄 64.0 岁), 2 414 名 (39.9%) 女性, 平均 HbA1c 7.2%, 糖尿病中位病程 6.3 年, 59.0% 的受试者接受了二甲双胍单药治疗。中位随访时间为 6.3 年。利格列汀组有 356/3 023 (11.8%) 发生了主要结局事件, 在格列美脲组有 362/3 010 (12.0%) 发生 ($HR=0.98$, $95\%CI:0.84\sim 1.14$, 非劣效性 $P<0.001$), 符合非劣效性标准,但不具有优效性。

模拟 CAROLINA 试验的 RWE 队列研究 (NCT03648424) 是利用美国两个商业健康保险索赔数据库 (Optum Clinformatics 和 IBM MarketScan) 和 Medicare 数据库来预测 CAROLINA 在真实世界中的结果^[23],并在 CAROLINA 的 RCT 完成之前对结果进行预测。该研究选择利格列汀和格列美脲初次使用者,以便得到一个能与 RCT 中试验人群可比较的研究群体。研究对象进入队列的日期定义为其医疗处方中第一次出现利格列汀或格列美脲的时间,在这个日期前个体需至少有连续 6 个月的医疗记录以确定新用药者。研究采用倾向评分 1:1 匹配法对 120 项基线特征进行了控制。根据 CAROLINA 试验的主要研究结局,对患者进行了随访。随访从进入队列后的第 1 天开始,直到停止治疗或更换到对照组、发生了预定结局事件、疗养院入院、退出队列或研究时间结束 (以最早发生事件为准)。如果治疗中断,研究会将暴露效应窗口延长至最后一个用药记录的后 30 d。考虑到常规护理中治疗中断的高发生率,研究选择了实际治疗分析 (as treated analysis, AT) 而非意向性治疗 (intention-to-treat, ITT),以避免 RWE 研究中经常出

现的暴露错分偏倚。该研究队列最终纳入 24 131 对 T2DM 患者,与格列美脲组相比,利格列汀组的 3P MACE 结局的风险无显著相关($HR=0.91, 95\%CI: 0.79\sim 1.05$),此结果符合 CAROLINA 试验的非劣效性假设。

CAROLINA RCT 与 RWE 研究结果比较发现,RWE 研究得到了与预期 RCT 相似的研究结论。这个项目的完成显示了 RWE 相较于 RCT 具有研究周期短和人力物力成本低两个显著的优势。CAROLINA RCT 花费近 8 年的时间才完成,除了需要医务人员和研究团队提供大量资源外,还需要招募大量符合标准的患者。相比之下,RWE 研究仅需一个由流行病学家和医学专家组成的小组,利用已有的 RWD 以及一个 RWD 分析平台就可以开展,且一般只需要几个月的时间就能得到研究结果。

在使用电子索赔数据开展 RWE 研究中,CAROLINA 的 RWE 研究项目组实现了足够的统计能力,并进行了可靠的混杂因素控制,研究发现在公布 RCT 研究结果之前 RWE 研究可以准确预测 CAROLINA 试验的结果,这表明 RWE 具有较高的可信度。而且根据 RWE 研究结果,由于 CAROLINA 试验的参与者的平均年龄低于模拟研究的人群(64 岁 vs. 70 岁),所以对于 RCT,利格列汀的年龄相关性心血管益处在实际人群中可能会更明显。

二、经验与启示

1. 研究经验总结:RCT DUPLICATE 旨在为选定的 RCT 样本提供严格的证据,以及非干预性的 RWE 研究在何时以及如何得出与 RCT 相同的结论^[26]。即使在偏倚风险很低的情况下,一些 RWE 研究也可能无法产生与 RCT 完全一致的结果^[27]。对于结果有统计显著性的 RCT,无偏倚的 RWE 研究结果与之符合监管一致性的概率在 80%~90% 之间,但未发现显著性差异的 RCT 中实现监管一致性的概率则更低^[13]。

RCT 和 RWE 的研究结果证实,选择具有相似适应证的阳性对照的 RCT 进行模拟会提高 RWE 的有效性。然而,即使在阳性对照的情况下,也不能完全保证 RCT 和 RWE 研究结果的一致性。RCT DUPLICATE 项目组开展的 RWE 研究都是基于严格筛选后的研究样本,最大程度上使用 RWD 测量结局指标,并考虑了重要的尽可能多的混杂因素。目前项目组遴选的 RCT 远不能覆盖所有临床话题,未来仍需要更多的证据来帮助理解 RWE 的适用场

景,从而确定 RWE 是否可以预测在不同治疗领域 RCT 的研究结果。

RCT DUPLICATE 研究仍存在一定的局限性,许多 RCT 会存在一些模糊的排除标准,如“在研究者的判断中”“不太可能存活至少 5 年”。这些标准使得利用 RWD 开展 RWE 研究时模拟 RCT 试验人群变得困难,并且很难在 RWE 研究中转化为一个可测量的指标。RWE 研究也难以精准地模拟 RCT 的治疗策略,包括坚持用药、剂量规格和附加治疗等。

RCT 中的安慰剂对照是开展 RWE 研究的主要难点。项目组试图通过识别接受标准治疗的患者来模拟安慰剂对照。然而在 RWD 中很难精准定义标准治疗,并且人群在 RCT 和 RWD 中选择的具体治疗方法可能存在差异。此外,尽管项目组试图统一重要结局的定义,但在如何定义和识别结局等方面仍可能存在差异。例如 CAROLINA 试验比较利格列汀和格列美脲,评估了中度或重度低血糖的风险。但对于这一结局,RCT 可以根据研究目的获取到这类信息,而基于医疗索赔数据的 RWE 研究难于直接获取这类结局指标。项目组在 RWD 中将低血糖定义为住院或急诊诊断为低血糖,或许可以解释 RCT 和 RWE 研究关于中重度低血糖风险比估计值的差异($HR=0.15, 95\%CI: 0.08\sim 0.29$; $HR=0.42, 95\%CI: 0.32\sim 0.56$),而在心血管结局上结果呈现高度一致性($HR=0.98, 95\%CI: 0.84\sim 1.14$; $HR=0.91, 95\%CI: 0.79\sim 1.05$)。

尽管存在上述局限性,RWE 研究结果与 RCT 结果之间的差异并不会阻碍前者在监管方面的价值。因为 RCT 和 RWE 之间存在本质上的“效力-效果差距”,RWE 还有尚未挖掘的价值。项目初期目标可能是为了加快推荐使患者获益的干预措施,尽管需要额外考虑医疗索赔系统的特点和患者的偏好。更长远的目标是减少从研究证据到更广泛的临床应用之间的漫长等待,辅助药物审批和上市后不同特征患者人群的实际疗效和安全性监测等监管决策应用。

项目组制定的一致性度量标准简单直观,但单纯依赖于 P 值和 $95\%CI$ 可能会遗漏 RWE 研究是否成功的相关细节。即使每个 RWE 研究都是完全无偏倚的,也会由于抽样的变异性导致 30 个 RWE 研究中的一些结果与 RCT 不同^[28]。但是该项目首次评估了大量高质量的 RCT 的可重复性,并在研究初期设置了一系列规范化的统一标准和透明化研究

流程,实现了在虚拟平台上的数据公开和细节追踪^[29-30]。尽管该项目的规模尚小,无法回答所有 RWE 研究设计的相关问题,未来的工作可以通过模拟更多的 RCT 来为不同的临床问题提供新的见解。

2. 对基于中国人群开展 RCT 模拟研究的启发: RCT 为监管机构提供了治疗效果评估的参考标准,对于有的 RWE 研究无法模拟 RCT 的结果仅能表明无法在真实世界中模拟,并不意味着原始 RCT 的研究结果是无效的。根据已有经验利用可用的 RWD,约 4% 的补充批准药物适应证的 RCT 可被 RWE 研究模拟^[26],表明 RCT 目前的可重复性普遍较低。然而,RCT DUPLICATE 项目经验指出,若一些 RCT 的主要结局难以在 RWE 研究中定义,此时可以在医疗索赔数据库中获取相似的替代指标,由此增强 RCT 可重复性。此外,由于 RCT DUPLICATE 项目组预先制定的定义限制,发现许多 RCT 是无法被平台上的数据模拟的,因为这些 RCT 需要的实验室指标无法在医疗索赔数据库中获取,但可能在其他 RWD 中很容易获取,如电子健康病历(electronic health record, EHR),后续需要在更广的数据源中开展模拟 RCT 的 RWE 研究。

如果基于中国人群开展类似的 RWE 研究,考虑到国内外人群本质上的差异,应尽量在研究人群的选择上保持一致性,即尽量选择基于亚洲人群的 RCT。关于 RWD 的选择,尽管 RCT DUPLICATE 项目组计划在医疗索赔数据库中模拟每个 RCT 的重要研究细节,但在其实际实施阶段发现 RWD 可以获取的测量值与 RCT 还是有差距的。如系列一所述,研究有时无法完全模拟某些 RCT 的纳入排除标准,这些标准往往依赖于索赔数据库中无法获得的临床测量指标。此外,不同于 RCT,实际临床治疗中患者的依从性不受研究者控制,且可能出现更多的停药和短期用药情况。这也启发我们在开展 RWE 研究时要着重关注所选 RCT 的对照和结局的测量能否在现有的 RWD 中获得。

根据我国的基本国情、医疗卫生系统的特点和可供选择的 RCT 特征,要更灵活地选择 RWD。除了 RCT DUPLICATE 项目组采用的医疗索赔数据库,一些其他来源的 RWD,如 EHR 和患者登记信息,可能会提供不同的、更广泛的可模拟的 RCT 选择,因为这些数据源通常涵盖更广泛和详细的临床信息,有助于在开展 RWE 研究时调整混杂因素和定义关键变量。如果将研究扩展到这些 RWD 来

源,就可能对 RCT DUPLICATE 项目目前没有包括的肿瘤、传染病和神经病学等其他医学领域的 RWE 进行可信度评估^[31-32]。但是,这些数据库并非没有局限性,例如门诊用药的数据在药房配药数据库中比在 EHR 中更容易获取^[33]。此外,EHR 可能会高估未按处方填写患者的用药情况,且可能遗漏患者在医疗机构外的治疗,从而导致严重的偏倚^[34]。所以在开展 RWE 研究时,根据研究领域和研究目的,科学灵活地选择合适的 RWD 来源和待模拟的 RCT,广泛探索 RWE 研究和 RCT 之间的可重复性、差距及其潜在的影响因素,旨在促进 RWE 研究结果的真实性和可靠性。

此外,考虑到患者隐私和数据安全性问题,在基于中国人群开展 RWE 研究时,可以参考 RCT DUPLICATE 项目组在开展研究前的一系列措施,由具有权威性的监管机构或者部门牵头,联合大型医疗卫生相关数据库搭建虚拟数据分析平台(如 RCT DUPLICATE 项目中的 Aetion 平台),在保证数据安全性的基础上,实现数据的共享和统计分析,并且制定相关研究规范标准,保证研究的透明化和流程化,如在研究方案确定后统一进行研究注册和细节公开等,进一步增强 RWE 研究结果的可信度。

在过去的几年间,RWE 受到很多学者和决策机构的关注,当前用于新型冠状病毒肺炎治疗相关的 RWE 也呈爆炸式增长^[35]。越来越多使用单臂治疗研究和非随机研究为 FDA 药物审批和标签拟定提供证据^[36-37]。我们也希望能通过 RWE 为患者和医疗决策相关机构提供一些有意义的补充证据。2021 年 *BMJ* 上发表了一篇关于 FDA 加速审批的药物存在上市后无临床效果且长期未被取消上市资格的相关问题的文章^[38],在 RWE 研究快速发展的当下,给我们以警醒,RWE 研究的初衷是为了更好地解决 RCT 结果不能推广至更广泛的人群以及长时间高成本等实际问题,补充决策证据体系,加速药品审批流程。但是在开展相关 RWE 研究时,需首先确保研究的质量。随着 RWE 研究的数据来源和研究方法的不断成熟,预计还会有更多的 RWE 研究涌现,将 RWE 纳入监管决策的机会也会越来越多。从长远来看,RWE 用于补充和提供决策证据依然是一个可行的方案,今后在开展 RWE 相关研究时,需要进一步完善相关规范并严格遵守标准,确保研究流程的透明化和高质量实施,切实为患者和相关决策者提供有效信息。

利益冲突 所有作者声明无利益冲突

作者贡献声明 石舒原:文献整合、论文撰写;刘佐相:文献收集;赵厚宇、聂晓璐:论文修订;韩晟、符祝、宋海波:论文修订、研究材料支持;姚晨、詹思延:指导性支持;孙凤:提出研究选题、获取研究经费

参 考 文 献

- [1] Bothwell LE, Greene JA, Podolsky SH, et al. Assessing the gold standard—lessons from the history of RCTs[J]. *N Engl J Med*, 2016, 374(22):2175-2181. DOI: 10.1056/NEJMms1604593.
- [2] Andre EB, Reynolds R, Caubel P, et al. Trial designs using real-world data: the changing landscape of the regulatory approval process[J]. *Pharmacoepidemiol Drug Saf*, 2020, 29(10):1201-1212. DOI:10.1002/pds.4932.
- [3] Corrigan-Curay J, Sacks L, Woodcock J. Real-world evidence and real-world data for evaluating drug safety and effectiveness[J]. *JAMA*, 2018, 320(9):867-868. DOI: 10.1001/jama.2018.10136.
- [4] Khozin S, Blumenthal GM, Pazdur R. Real-world data for clinical evidence generation in oncology[J]. *J Natl Cancer Inst*, 2017, 109(11):dix187. DOI:10.1093/jnci/dix187.
- [5] 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. DOI: 10.1056/NEJMsb1609216.
- [6] Bonamici S. H. R. 34-21st Century cures act[EB/OL]. [2022-03-15]. <https://www.congress.gov/bill/114th-congress/house-bill/34>.
- [7] U.S. Food and Drug Administration (FDA). Framework for FDA's real-world evidence program (2018) [EB/OL]. [2022-03-15]. <https://www.fda.gov/media/120060/download>.
- [8] European Medicines Agency (EMA). EMA regulatory science to 2025. Strategic reflection[EB/OL]. [2022-03-15]. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection_en.pdf.
- [9] Chen PY. Key considerations in using real-world evidence to support drug development[EB/OL]. [2020-09-23][2022-03-15]. <https://www.amstat.org/meetings/biop/2020/onlineprogram/ViewPresentation.cfm?file=301264.pdf>.
- [10] 石舒原,周庆欣,孙凤,等.真实世界证据与随机对照试验:RCT DUPLICATE项目成果[J].*药物流行病学杂志*, 2019, 28(11):757-762. DOI:CNKI:SUN:YWLX.0.2019-11-013.
- [11] Shi SY, Zhou QX, Sun F, et al. Real-world evidence and randomized controlled trials: the outcome of RCT DUPLICATE[J]. *Chin J Pharmacoepidemiol*, 2019, 28(11):757-762. DOI:CNKI:SUN:YWLX.0.2019-11-013.
- [12] 姚晓莹,张靖雪,詹思延.真实世界证据与随机对照试验:RCT DUPLICATE项目概述[J].*药物流行病学杂志*, 2019, 28(8):495-497, 517.
- [13] Yao XY, Zhang JX, Zhan SY. Real-world evidence and randomized controlled trials: the overview of RCT DUPLICATE[J]. *Chin J Pharmacoepidemiol*, 2019, 28(8):495-497, 517.
- [14] 石舒原,赵厚宇,周庆欣,等.真实世界证据与随机对照试验:RCT DUPLICATE项目方法学介绍[J].*药物流行病学杂志*, 2020, 29(3):198-205. DOI:CNKI:SUN:YWLX.0.2020-03-012.
- [15] Shi SY, Zhao HY, Zhou QX, et al. Real-world Evidence and Randomized Controlled Trials: the Methodology of RCT DUPLICATE[J]. *Chin J Pharmacoepidemiol*, 2020, 29(3):198-205.
- [16] Franklin JM, Paterno E, Desai RJ, et al. Emulating randomized clinical trials with nonrandomized real-world evidence studies: first results from the RCT DUPLICATE initiative[J]. *Circulation*, 2021, 143(10):1002-1013. DOI:10.1161/circulationaha.120.051718.
- [17] Marso SP, Daniels GH, Brown-Frandsen K, et al. Liraglutide and cardiovascular outcomes in type 2 diabetes[J]. *N Engl J Med*, 2016, 375(4):311-322. DOI: 10.1056/NEJMoa1603827.
- [18] Wiviott SD, Raz I, Bonaca MP, et al. Dapagliflozin and cardiovascular outcomes in type 2 diabetes[J]. *N Engl J Med*, 2019, 380(4):347-357. DOI:10.1056/NEJMoa1812389.
- [19] Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes [J]. *N Engl J Med*, 2015, 373(22):2117-2128. DOI:10.1056/NEJMoa1504720.
- [20] Neal B, Perkovic V, Mahaffey KW, et al. Canagliflozin and cardiovascular and renal events in type 2 diabetes[J]. *N Engl J Med*, 2017, 377(7):644-657. DOI:10.1056/NEJMoa1611925.
- [21] Paterno E, Goldfine AB, Schneeweiss S, et al. Cardiovascular outcomes associated with canagliflozin versus other non-gliiflozin antidiabetic drugs: population based cohort study[J]. *BMJ*, 2018, 360:k119. DOI:10.1136/bmj.k119.
- [22] Rosenstock J, Perkovic V, Johansen OE, et al. Effect of Linagliptin vs. placebo on major cardiovascular events in adults with type 2 diabetes and high cardiovascular and renal risk: the CARMELINA randomized clinical trial[J]. *JAMA*, 2019, 321(1):69-79. DOI:10.1001/jama.2018.18269.
- [23] Green JB, Bethel MA, Armstrong PW, et al. Effect of Sitagliptin on cardiovascular outcomes in type 2 diabetes [J]. *N Engl J Med*, 2015, 373(3):232-242. DOI: 10.1056/NEJMoa1501352.
- [24] Scirica BM, Bhatt DL, Braunwald E, et al. Saxagliptin and cardiovascular outcomes in patients with type 2 diabetes mellitus[J]. *N Engl J Med*, 2013, 369(14):1317-1326. DOI: 10.1056/NEJMoa1307684.
- [25] Rosenstock J, Kahn SE, Johansen OE, et al. Effect of Linagliptin vs. glimepiride on major adverse cardiovascular outcomes in patients with type 2 diabetes: the CAROLINA randomized clinical trial[J]. *JAMA*, 2019, 322(12):1155-1166. DOI:10.1001/jama.2019.13772.
- [26] Paterno E, Schneeweiss S, Gopalakrishnan C, et al. Using real-world data to predict findings of an ongoing phase IV cardiovascular outcome trial: cardiovascular safety of Linagliptin versus glimepiride[J]. *Diabetes Care*, 2019, 42(12):2204-2210. DOI:10.2337/dc19-0069.
- [27] Wallentin L, Becker RC, Budaj A, et al. Ticagrelor versus clopidogrel in patients with acute coronary syndromes[J]. *N Engl J Med*, 2009, 361(11):1045-1057. DOI: 10.1056/NEJMoa0904327.
- [28] Wiviott SD, Braunwald E, McCabe CH, et al. Prasugrel versus clopidogrel in patients with acute coronary syndromes[J]. *N Engl J Med*, 2007, 357(20):2001-2015. DOI:10.1056/NEJMoa0706482.
- [29] Franklin JM, Pawar A, Martin D, et al. Nonrandomized real-world evidence to support regulatory decision making: process for a randomized trial replication project [J]. *Clin Pharmacol Ther*, 2020, 107(4):817-826. DOI: 10.1002/cpt.1633.
- [30] Franklin JM, Dejene S, Huybrechts KF, et al. A bias in the evaluation of bias comparing randomized trials with nonexperimental studies[J]. *Epidemiol Methods*, 2017, 6(1):20160018. DOI:10.1515/em-2016-0018.
- [31] Thompson D. Replication of randomized, controlled trials using real-world data: what could go wrong? [J]. *Value Health*, 2021, 24(1):112-115. DOI: 10.1016/j.jval.2020.09.015.
- [32] Wang SV, Verpillat P, Rassen JA, et al. Transparency and reproducibility of observational cohort studies using large healthcare databases[J]. *Clin Pharmacol Ther*, 2016, 99(3):325-332. DOI:10.1002/cpt.329.
- [33] Kim SC, Solomon DH, Rogers JR, et al. Cardiovascular safety of Tocilizumab versus tumor necrosis factor inhibitors in patients with rheumatoid arthritis: a multi-database cohort study[J]. *Arthritis Rheumatol*, 2017, 69(6):1154-1164. DOI:10.1002/art.40084.
- [34] Urquhart L. FDA new drug approvals in Q1 2019[J]. *Nat Rev Drug Discov*, 2019. DOI: 10.1038/d41573-019-00070-3.
- [35] Mullard A. 2017 FDA drug approvals[J]. *Nat Rev Drug Discov*, 2018, 17(2):81-85. DOI:10.1038/nrd.2018.4.
- [36] Steiner JF, Prochazka AV. The assessment of refill compliance using pharmacy records: methods, validity, and applications[J]. *J Clin Epidemiol*, 1997, 50(1):105-116. DOI:10.1016/s0895-4356(96)00268-5.
- [37] Lin KJ, Glynn RJ, Singer DE, et al. Out-of-system care and recording of patient characteristics critical for comparative effectiveness research[J]. *Epidemiology*, 2018, 29(3):356-363. DOI: 10.1097/ede.0000000000000794.
- [38] Pundi K, Perino AC, Harrington RA, et al. Characteristics and strength of evidence of COVID-19 studies registered on ClinicalTrials.gov[J]. *JAMA Intern Med*, 2020, 180(10):1398-1400. DOI:10.1001/jamainternmed.2020.2904.
- [39] Hatswell AJ, Baio G, Berlin JA, et al. Regulatory approval of pharmaceuticals without a randomised controlled study: analysis of EMA and FDA approvals 1999-2014[J]. *BMJ Open*, 2016, 6(6):e011666. DOI:10.1136/bmjopen-2016-011666.
- [40] Califf RM, Hernandez AF, Landray M. Weighing the benefits and risks of proliferating observational treatment assessments: observational cacophony, randomized harmony[J]. *JAMA*, 2020, 324(7):625-626. DOI:10.1001/jama.2020.13319.
- [41] Mahase E. FDA allows drugs without proven clinical benefit to languish for years on accelerated pathway[J]. *BMJ*, 2021, 374:n1898. DOI:10.1136/bmj.n1898.