

·循证医学理论与方法研究·

# 真实世界研究的方法与实践

黄卓山, 罗艳婷, 刘金来

(中山大学附属第三医院心内科, 广州 510630)

[摘要] 随机对照临床试验是目前获得证据、制定诊疗策略的主要依据之一,但是,随机对照试验并不总能代表临床实践中的真实情况。真实世界研究是观察性研究,通过“真实世界样本”来反映真实世界总体。临床试验关注效力研究,而真实世界研究则关注效果研究。真实世界研究需要使用多重倾向性评分来减少协变量在组间分布的不均衡性。近年来,真实世界研究正越来越多地应用到临床研究中,可将经验医学与循证医学结合起来。临床试验和真实世界研究是承启关系,前者的结果需要后者的进一步验证及拓展,二者综合考虑才是最佳选择。

[关键词] 真实世界研究;随机对照临床试验;倾向性评分;效果研究

[中图分类号] R195.1;R544.1 [文献标识码] A DOI:10.3969/j.issn.1671-5144.2014.06.014

## Method and Practice of Real World Study

HUANG Zhuo-shan, LUO Yan-ting, LIU Jin-lai

(Division of Cardiovascular Medicine, The Third Affiliated Hospital of Sun Yat-sen University, Guangzhou 510630, China)

**Abstract:** Randomized controlled trial (RCT) is so far considered as an important way to provide evidence to make up therapeutic strategy. However, sometimes it is limited to represent the authentic condition during clinical practice. Whereas, Real world study (RWS), a kind of observational studies, is applied to present the real world population through analysis of “real world sample”. RCT pays more attention to efficacy trials. While RWS more concentrates on effectiveness trials. RWS requires giving propensity scores to decrease covariant malconformation among groups. In the past few years, RWS has been increasingly applied in clinical studies to integrate experience medicine with evidence-based medicine. Results of RCT need to be tested and expanded forward in real world study. Therefore, comprehensive evaluations by both RCT and RWS are supposed to be the best selection.

**Key words:** real world study; randomized controlled trial; propensity score; effectiveness trials

### 1 从大型研究结果引出的思考

Symlicity HTN-1 是首个肾交感神经消融 (renal sympathetic denervation, RSD) 治疗高血压的临床研究<sup>[1]</sup>, 45 例难治性高血压患者接受了 RSD 治疗, 结果显示术后即刻与术前相比平均诊室血压降低, 术后 1 个月显著降低。随后, 该研究进一步扩

大样本量至 153 例, 随访 2 年结果显示, 患者血压水平有明显的下降<sup>[2]</sup>。最终 3 年的随访结果显示<sup>[3]</sup>, 术后 12 个月、24 个月、36 个月, 患者血压分别降低 27/12 mmHg、30/13 mmHg、32/14 mmHg。术后随访 1 个月、6 个月、12 个月、24 个月、36 个月时, 分别有 69%、81%、83%、85%、93% 的患者收缩压降低 10 mmHg 以上, 提示部分患者可能存在延迟起效。虽然该研究提示 RSD 具有良好的安全性和有效性, 但是不能排除安慰剂效应和观察者偏倚, 因此提出需要进行前瞻性随机临床研究。

为解决 HTN-1 试验设计的局限性, 2010 年 Symlicity HTN-2 试验采用多中心、前瞻性、随机

[作者简介] 黄卓山 (1988-), 男, 广州人, 硕士研究生, 从事冠心病的基础与临床研究。

[通讯作者] 刘金来, Tel: 020-85252168; E-mail: lj.lai@medmail.com.cn

对照设计<sup>[4]</sup>, 106例难治性高血压患者按1:1比例随机分组, 一组接受导管消融术和常规降压药物治疗, 另一组仅接受降压药物治疗。6个月随访时, 与基线血压水平相比, 手术治疗组血压下降32/12 mmHg, 药物治疗组血压无明显改变, 两组间平均血压差为33/11 mmHg, 存在明显差异。随后, 该试验对照组中35例患者接受了RSD治疗, 再随访6个月后平均收缩压显著下降(24 mmHg)<sup>[5]</sup>。此项随机试验证实了以前的研究结果, 但也同样有局限性。第一, 样本量相对较小。第二, 随访时间有限。第三, 继发性高血压筛查有限, 患者用药依从性不可控。第四, 以诊室血压作为主要观察结果而非24小时动态血压监测易产生测量误差和观察者偏倚。第五, 顽固性高血压定义和药物治疗不严格。第六, 患者不设盲故不能排除安慰剂效应。

基于以上理由, Symplicity HTN-3多中心试验于2011年开展<sup>[6]</sup>。该研究为单盲、随机、对照研究, 入选530名难治性高血压患者, 随机分为RSD组和单纯药物治疗组(按2:1比例随机分组), 主要研究终点指标为术后6个月患者诊室收缩压变化, 次要终点指标是术后6个月患者24小时血压动态监测的平均变化, 主要安全性终点指标是术后6个月内主要不良事件发生率。该试验设有非常严格的入组标准, 排除首次动态血压评估平均收缩压<135 mmHg的患者, 同时, 对临床观察人员设盲并应用自动血压计监测诊室血压。此外, 所有随机进入对照组的患者均需接受假手术治疗。然而, 出乎意料的是, 试验没有达到最初设定的主要终点。尽管没有安全性方面的相关事件发生, 但受试者的收缩压未曾持续性下降。

这是一个让人意外的结果, 但底特律医学中心医学博士Cindy Grines说:“我认为这个研究结果并不能终结这种去肾神经支配手术的开展。相反这个试验结果可能仅仅是由于方法学的错误所造成”。随机对照临床试验(randomized controlled trial, RCT)是目前公认的治疗性或预防性临床研究方法的金标准, 用于临床医学研究已有50多年的历史。RCT是将符合要求的研究对象分别分配到试验组和对照组, 在同一条件下, 同期按预期方案给予干预措施和观察, 对试验结果进行测定和评价。RCT的严谨设计可以避免研究中可能存在的选择性偏倚和测量性偏倚, 确保研究结果真实性, 是获得证据、制定诊疗策略的主要依据之一。RCT研究获得的只是理想条件下或某一年龄段或疾病亚组人

群干预措施的结果, 尚不能代表临床实践中的真实状况。此外, RCT常涉及到医学伦理学问题; 观察时间很有限, 有时不可避免地出现过度的失访、退出、沾染或干扰影响研究结果事件; 不能满足不同危险度个体的循证治疗; 甚至某些疾病无法实施RCT(如手术治疗与非手术治疗比较)。

那么我们是否会有更好的选择来进行临床研究呢?

## 2 随机对照试验和真实世界研究

真实世界研究(real world study, RWS)起源于实用性临床试验, 是指在较大的样本量(覆盖具有代表性的更大受试人群)基础上, 根据患者的实际病情和意愿非随机选择治疗措施, 开展长期评价, 并注重有意义的结局治疗, 以进一步评价干预措施的外部有效性和安全性, 1993年Kaplan等<sup>[7]</sup>首次在论文中使用。其涵盖的范围较随机对照试验更宽, 除治疗性研究外, 还可用于诊断、预后、病因等方面的研究。

针对严格的解释性RCT纳入人群限制较多用药条件控制严格, 使得研究结果的内部真实性较高, 外部真实性却较差, 研究结果的实际应用推广受限等问题, 国际上在实用性RCT的基础上, 提出了真实世界研究的概念和方法, 通过“真实世界样本”来反映真实世界总体, 最初主要用于对药物临床不良反应的监测, 就某药物在现实临床中监测到的不良反应, 采用药物流行病学分析方法, 辨别是否属于该药的不良反应; 其后逐步发展到上市药物有效性和安全性再评价及临床干预措施的评价, 主要是在不限定临床干预措施的情况下研究其效果。RCT从医疗者角度评估医疗手段的“效力”(efficacy), 是确定医疗手段有效性和安全性的标准方法。RWS在真实医疗过程中进行观察性临床研究, 用于观察药物、新疗法在广泛真实医疗过程中的疗效和不良反应, 是从患者角度评估医疗措施的“效果”(effectiveness)。需要注意的是, 效力并不等于效果。RCT关注效力研究(efficacy trials), 即药物与干预措施能否在理想、严格控制的环境下产生预期的效果, 着重于内部有效性, 不易普遍化; 选择人群观察时间短, 人群样本小, 测量手段主要有中间终点如糖化血红蛋白、肺功能-第一秒用力呼气量、血压等, 当然如果随访时间较长也有硬终点如病死率等。RWS关注效果研究(effectiveness trials), 即评价药物在真实临床环境下的治疗效果, 重在外部的有效性; 缺乏控制, 存在选择

性偏倚、观察性偏倚等混杂因素, 需要有不同于 RCT 的统计方法进行校正, 测量手段有死亡率、无症状时间、患者生存质量等<sup>[8-9]</sup>。

真实世界数据或真实世界研究覆盖多种研究类型及数据资源, 包括患者注册研究、已有的电子健康记录、常规收集的服药数据、患者原始数据、人群健康调查等, 数据来源广泛。在多种数据类型中, 临床终点指标(如发病率、死亡率)、患者报告的终点事件、安全性数据、疾病进展的自然病史均可直接供他国参考, 而处方类型、经济学模型数据、流行病学数据则要分别结合当地经济及政策、疾病及国家具体情况考虑, 治疗路径、使用资源、患者服务经验则因各国具体国情不一无法参考。

在具体的设计和各自的局限性上, RCT 和

RWS 有很大的差别(见表 1)。由于 RWS 是观察性研究, 研究对象纳入限制较少、人群异质性较大、自主选择治疗措施可造成的选择偏倚以及重要预后因素在组间分布的不均衡性, 需要进行多重倾向性评分减少这些影响。倾向性评分的概念最早由 Rosenbaum 和 Rubin 在 1983 年首次提出<sup>[10]</sup>, 其基本原理是用一个倾向指数表示多个协变量的影响, 根据倾向指数进行不同对比组间的匹配、分层或加权, 即均衡对比组间协变量的分布, 最后在协变量分布均衡的层内或者匹配组中估计处理效应。在大样本的情况下, 经过倾向评分值调整的组间个体, 除了暴露因素和结局变量分布不同外, 其他协变量应当均衡可比, 相当于进行了“事后随机化”, 使观察性数据达到“接近随机分配数据”的效果<sup>[11]</sup>。

表 1 RCT 与 RWS 设计的区别和局限性

指标	RCT	RWS
患病人群	严格有限样本, 设计者制定纳入标准和排除标准。	宽泛大样本, 无需制定纳入标准和排除标准, 符合治疗适应证的患者均可纳入。
分组方法	按随机、安慰剂对照的原则将样本分为治疗组及安慰剂对照组。	在非随机开放、无安慰剂对照的情况下将患者分为暴露组和公认有效的对照组。
研究过程	在较短时间内通过研究方案的治疗和随访得出结果。	通过长时间专门的治疗和随访(质控伦理), 在完备注册信息和数据库支持下得出结果。
局限性	结论外推性差, 这一直是制约其发展的重要原因。	观察者偏倚, 且样本量大、随访时间长造成其研究成本高, 庞大的数据收集量也增大了工作难度, 并有可能存在潜在编码错误和数据丢失的问题。

虽然二者差别巨大, 但并不是对立关系, 而是互补与承启关系。RCT 是评价任何临床干预措施的基础, 用于评价有效性和安全性, 没有 RCT, 任何外部有效性的结果都将受到质疑。在 RCT 的基础上制定相应的治疗指南, 新的临床干预措施得以真正用于临床, 但指南是一种推荐, 它告诉医生哪些应该做或可以做, 而不是哪些必须做, 指南不能替代临床经验。所以, 就需要 RWS 作为有效补充, RWS 用于决定效应性, 能够用于决定临床实践中真实的效益、风险和治疗的临床价值, 使临床研究的结论在 RCT 后回归真实世界。

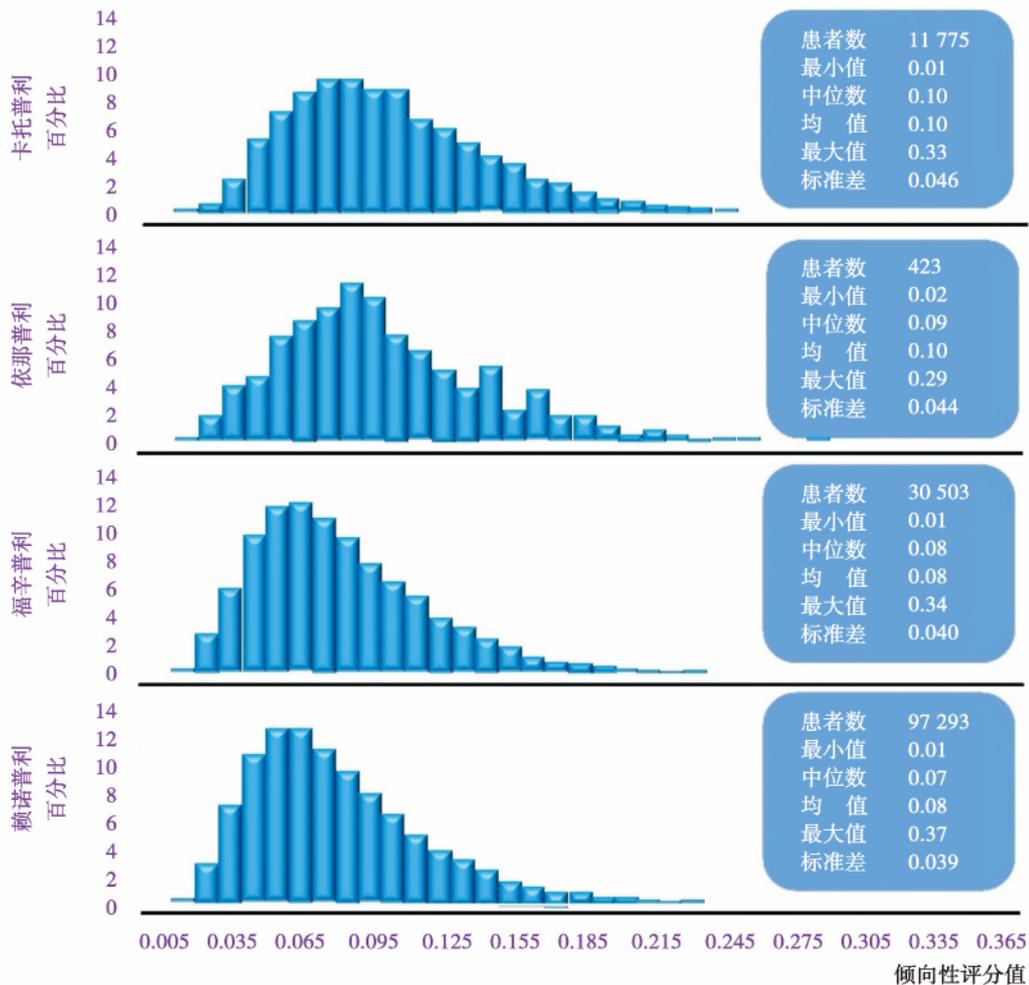
### 3 真实世界研究例子

2012 年, Chitnis 等用多重倾向性分析的方法比较几种血管紧张素转换酶抑制剂(angiotensin-converting enzyme inhibitors, ACEI)对心力衰竭患者死亡率的影响<sup>[12]</sup>, 这是真实世界研究的一个例

子。该研究纳入了 139 994 例心力衰竭患者, 其中 69.50%(97 293 例)接受赖诺普利、21.79%(30 503 例)接受福辛普利、8.41%(11 775 例)接受卡托普利和 0.30%(423 例)接受依那普利的治疗, 以卡托普利为参照, 比较不同 ACEI 治疗的粗略死亡率。研究中共有 47 个基线危险因素(协变量), 在这些基线危险因素中仅 11 项治疗组间无差异, 36 项差异有统计学意义, 但是建立倾向性评分模型、经过标准化倾向性权重评估后, 绝大多数协变量差异无统计学意义, 仅 3 项(肾衰竭、消化系统疾病、贫血和凝血异常)仍有差异, 倾向性评分后显示各治疗组间具有可比性(图 1)。最终研究结论显示, 与卡托普利相比, 福辛普利和赖诺普利显著降低心力衰竭患者死亡风险, 依那普利与卡托普利在此方面无显著差异。该研究首次在现实世界人群证实了不同 ACEI 对于心力衰竭患者预后会产生不同的影响, 对未来 ACEI 的治疗及研究提

供了新的参考思路,可以使更多患者受益于更好 ACEI 的治疗。同时,多重倾向性分析方法的应用,

为现实世界治疗数据分析提供了有效的方法和思路。



各组之间的重叠允许对治疗和对照药物的效应进行直接估计,体现了对所有观测的背景特征差异的校正。各组间的重叠相当大,显示各组之间具有可比性。

图 1 倾向性评分柱状图

2011年,Petrella等在《临床治疗学》杂志发表血管紧张素受体拮抗剂(angiotensin receptor blocker, ARB)对比其它类降压药物真实世界研究回顾性分析<sup>[13]</sup>,回顾性分析纳入17万加拿大高血压患者数据,坚持使用降压药治疗超过9个月,旨在比较ARB对比非ARB治疗方案的降压疗效,主要观察终点指标为血压达标率(目标血压值<140/90 mmHg)。真实世界临床数据显示,以ARB为基础的联合治疗方案血压达标率更高,其中,ARB/氢氯噻嗪联合治疗方案血压达标率较高。

2011年,Vegter等进行ACEI和ARB的依从性、坚持率及换药模式的研究<sup>[14]</sup>,纳入1999–2010年间的超过50 000例患者处方,处方用药包括

ACEI:卡托普利、依那普利、赖诺普利、培哚普利、雷米普利、福辛普利;ARB:氯沙坦、缬沙坦、厄贝沙坦、坎地沙坦、奥美沙坦,用药模式包括:剂量、依从性、坚持率、换药。研究结果显示,ACEI与ARB之间的依从性无差别;治疗3年后,ACEI与ARB之间的坚持率无显著差别;福辛普利的依从性高于培哚普利,且显著高于依那普利;福辛普利的坚持率高于培哚普利。

通过以上的3个例子,我们可以看到RWS的重要作用,作为试验与实践的两个组成部分,RCT与RWS相互补充。高血压领域“真实世界研究”类文献逐年递增,通过PubMed,以hypertension + real-world为关键词搜索高血压领域在2003–

2012年之间发表的实效研究相关文献资料结果(图2)表明,真实世界研究正越来越多地应用到临床研究中。计算机技术支持的临床实践所产生的大量临床数据为进行科学研究奠定了基础。在大数据背景下,可基于现有数据库中的数据进行分析,来支持不同种类的业务,如医疗费用及报销、患者病史分析、归档影像分析、实时临床数据分析,并可应用于远程会诊,无线移动技术医疗信息采集、分析及共享,以支持医疗协同、临床决策支持和公共卫生管理。过去的医学模式是经验医学,强调的是个人经验与实践,随着RCT的发展,现在是循证医学的年代,RWS则可将经验医学与循证医学结合起来。

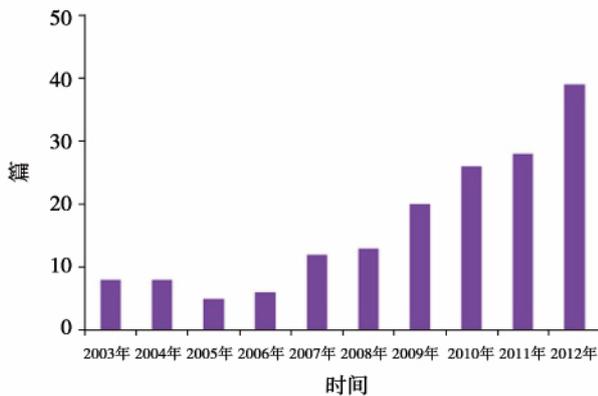


图2 2003-2012年高血压领域发表的RWS相关文献资料

#### 4 结 语

毋庸置疑,21世纪是循证医学时代,是数据引导策略的新纪元!这就意味着,大样本数据库所能带给我们的绝不仅仅是数字的堆砌,其真正价值在于能帮助我们深入、全面和准确地研究或解决重大临床问题,使结论更接近于“真实世界”,进而优化或革新临床实践。RCT用于评价效力,而RWS用于决定效果,用于决定临床实践中真实的效益、风险和临床价值。RCT提供证据,荟萃分析确定证据,RWS验证证据,临床治疗经验个体化应用证据,这一切源自临床实践,回答临床问题,总结治疗推荐,回归实践检验。RCT与RWS对同一个问题的论证是承启关系,RCT结果需要RWS的进一步验证及拓展,二者综合考虑才是最佳选择。RWS是开展治疗措施再评价研究的一种新思路,可进一步保障药物的有效性和安全性。但是,如何解决RWS中观察者偏倚、成本高等问题,仍然需要进一步继续完善。

#### [参 考 文 献]

- [1] Krum H, Schlaich M, Whitbourn R, et al. Catheter-based renal sympathetic denervation for resistant hypertension: A multicentre safety and proof-of-principle cohort study [J]. *Lancet*, 2009,373(9671):1275-1281.
- [2] Symplicity HTN-1 Investigators. Catheter-based renal sympathetic denervation for resistant hypertension: Durability of blood pressure reduction out to 24 months [J]. *Hypertension*, 2011,57(5):911-917.
- [3] Krum H, Schlaich MP, Sobotka PA, et al. Percutaneous renal denervation in patients with treatment-resistant hypertension: Final 3-year report of the Symplicity HTN-1 study [J]. *Lancet*, 2014,383(9917):622-629.
- [4] Symplicity HTN-2 Investigators, Esler MD, Krum H, Sobotka PA, et al. Renal sympathetic denervation in patients with treatment-resistant hypertension (The Symplicity HTN-2 Trial): A randomized controlled trial [J]. *Lancet*, 2010,376(9756):1903-1909.
- [5] Esler MD, Krum H, Schlaich M, et al. Renal sympathetic denervation for treatment of drug-resistant hypertension: One-year results from the Symplicity HTN-2 randomized, controlled trial [J]. *Circulation*, 2012,126(25):2976-2982.
- [6] Bhatt DL, Kandzari DE, O'Neill WW, et al. A controlled trial of renal denervation for resistant hypertension [J]. *N Engl J Med*, 2014,370(15):1393-1401.
- [7] Kaplan NM, Sproul LE, Mulcahy WS. Large prospective study of ramipril in patients with hypertension. CARE Investigators. [J]. *Clin Ther*, 1993,15(5):810-818.
- [8] 李敏,时景璞,于慧会. 真实世界研究与随机对照试验、单病例随机对照试验在临床治疗性研究中的关系比较 [J]. *中华流行病学杂志*, 2012,33(3):342-345.
- [9] 李敏,时景璞. 疗效比较研究的方法学应用及其实施过程 [J]. *中华流行病学杂志*, 2012,33(11):1184-1188.
- [10] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects [J]. *Biometrika*, 1983,70(1):41-55.
- [11] 王永吉,蔡宏伟,夏结来,等. 倾向指数的基本概念和研究步骤 [J]. *中华流行病学杂志*, 2010,31(3):347-348.
- [12] Chitnis AS, Aparasu RR, Chen H, et al. Effect of certain angiotensin-converting enzyme inhibitors on mortality in heart failure: A multiple-propensity analysis [J]. *Res Social Adm Pharm*, 2012,8(2):145-156.
- [13] Petrella R, Michailidis P. Retrospective analysis of real-world efficacy of angiotensin receptor blockers versus other classes of antihypertensive agents in blood pressure management [J]. *Clin Ther*, 2011,33(9):1190-1203.
- [14] Vegter S, Nguyen NH, Visser ST, et al. Compliance, persistence, and switching patterns for ACE inhibitors and ARBs [J]. *Am J Manag Care*, 2011,17(9):609-616.

[收稿日期] 2014-09-18