

文章编号: 1672-3384 (2008) -06-0030-03

循证医学和循证药学新进展

【作者】 詹思延

北京大学循证医学中心 (北京 100083)

【中图分类号】 R95

【文献标识码】 B

循证医学强调证据分级, 在评价药物疗效时“最佳证据”主要来自随机对照试验 (randomized controlled trial, RCT) 及其系统评价和 meta 分析。经典的 RCT 通常要求研究对象患单一疾病, 采用标准治疗和单一干预措施, 从而评价干预措施在理想状态下的特异性功效或效力 (efficacy)。在临床实际中, 病人经常罹患多种疾病, 同时接受多种治疗措施, 最终的疗效是欲研究的干预措施与其他各种处理因素 (如治疗方式、管理、辅助治疗等) 的综合效果。这样一来, RCT 作为在理想条件下开展的理论疗效研究, 难以提供临床实际疗效或称叫效果 (effectiveness) 的证据 (见表 1); 此外, RCT 常常需要制定严格的人选排除标准, 导致研究对象代表性不够, 进而不同程度地限制了其结论推论到总体; 第三, 盲法有时难以实现, 尤其对医疗器械进行评价时。因此, RCT 在评价常规医疗实践中不同治疗方法对病人的实际疗效时面临一定的挑战^[1,2]。

表 1 理论疗效研究和实际疗效研究的特征

理论疗效研究 (efficacy)	实际疗效研究 (effectiveness)
临床试验——理想条件下	临床实践——实际条件下
治疗组 VS 安慰剂组	多种治疗选择, 互相比较
患单一疾病的病人	患多种疾病的病人 (通常在功效研究中被排除)
实际使用人群可能会被排除 (如老年人)	一般不对研究对象做任何限制
通过测量替代终点、生物标志等评估短期效果	通过测量临床状况改善、生活质量、残疾、死亡等指标评估长期效果

要进行实际的疗效比较就需要在更大范围内观察治疗措施的有效性和安全性。但实施 RCT 通

常要有很高的投入, 当经典 RCT 因设计上的一些要求或资金、人力不能满足而无法开展的情况下, 应该考虑其他的研究设计类型。本文仅就近年来国外学者提出并尝试的几种新方法进行概括介绍。

大规模试验 (large trial, mega trial)

大规模试验适合用来评估疗效不是很大的干预措施^[3], 如抗高血压药可以将中风危险从 4% 降低到 2%, 就需要上万人 5 到 10 年的观察。大规模试验一般都是多国、多中心的研究, 它在设计和实施方面与一般试验存在很多不同 (见表 2)。大规模试验也常用于比较两个效果相当的治疗, 即等效试验 (equivalence trial)。

整群随机试验 (cluster randomized trial, CRT)

一般的 RCT 是以个体为随机分组单位, 整群试验则以群组, 如家庭、居委会、医院、工厂、村庄等为分组单位。在一个群组中的所有个体都接受相同的干预措施。与标准的 RCT 相比, 整群试验的统计效率比较低, 组间可比性差, 存在设计效应, 因此需要较大的样本量。但整群试验能解决特定的问题, 首先是评估那些只能针对整个人群施加的措施, 如卫生政策、行政和社会措施、管理模式、公共卫生措施等; 其次是评估在个体病人之间会发生“污染”的干预措施。所谓污染, 就是无干预的对照组一定程度上获得了治疗的干预, 如健康教育、医学筛查等, 研究者无法保证对照组不会从其他渠道接受健康教育或同样的筛查服务^[4]。

在疗效研究中, 如果以区域作为分组单位, 整群试验能够评估某治疗方法在真实临床实践环

表 2 一般临床试验与大规模临床试验的主要区别

特征	一般临床试验	大规模临床试验
研究目的	多属于理论疗效研究，即估计在理想条件下的效果	多属于临床疗效研究，即估计在常规治疗条件下的效果
样本量大小	几十或几百人	每组可以多达几万人
研究组织者数目	一般少于 10 人	非常多，有时可多达几百人
参与研究单位数	多只有 1 个单位	多国家、多城市、多中心
病人入选条件	多只限于 1 类病	可以入选各种不同类型的具有同一疾病的病人
治疗的安排	严格限制各比较组可能接受的其他治疗；用最好的医生、以最优的方式提供治疗	除研究的措施外，对各比较组可能接受的其他治疗不设特别限制；以实际治疗水平和条件提供治疗
随访的频率	随访频繁，强度高	常规治疗时的随访频率
病人依从性	采取特别措施提高依从性，病人依从性高	只提供常规治疗时采取的提高依从性的措施，病人依从性低
结果的诠释	若显示有效，实际中可能无效，应用时需慎重； 若显示无效，肯定无效	若显示有效，实际中一定有效；若显示无效，在理想治疗条件下可能会有效

(唐金陵等, 2007)

境中的实际效果，其结论不仅具有一般性，而且对于制定“真实世界”的决策（如临床实践指南、卫生政策等）非常有用。如果现有医疗卫生系统能够提供关于研究对象及其治疗方案、结局的信息时，实施整群试验可以节省大量的资金和时间^[5]。

3 实用性临床试验及开发以实践为基础的证据来改善临床实践

实用性临床试验（pragmatic clinical trials, PCTs）主要是针对解释性试验，即经典的 RCT 而言的，一般是类试验，主要用于常规临床实践中疗效对比研究^[6]。这种试验通常选择两种待比较的临床干预措施或方案；采用宽泛的入选标准，以保证结论能够最大程度地外推；观察多种结局，其中包括功能状态和医疗服务情况；研究在实际临床实践条件下进行，并尽可能地干预常规治疗。开发以实践为基础的证据（practice - based evidence for clinical practice improvement, PBE-CPI）为实施此类试验提供了可行的方法^[7]。与 RCT 不同，PBE-CPI 属于一种观察性研究，不需要改变现有的治疗方案就能够评价某种特定干预措施的效果。它通过掌握有关病人个体特征、医疗过程和临床结局等深入而复杂的信息，判定个体医疗过程对临床结局的影响，从而控制病人之间的差异。RCT 和 PBE-CPI 两种研究方法特点的比较见表 3。

在临床实践中，多种干预措施常常同时给予病人，各种处理措施之间的相互作用可能会对治疗效果产生很大影响。RCT 不可能对常规医疗工作中所有处理措施间的相互作用都进行验证，但医疗机构内部和各医疗机构之间如果对患病情况迥异的病人进行详尽的记录，则为联合干预的疗效比较研究提供了可能。PBE-CPI 就是充分利用医疗人员日常医疗活动的记录，以不加干预的方式分析处理措施的效果，这对于判定各种处理方法的相对疗效很重要。PBE-CPI 可用来评价当前的临床实践，并在分析各种临床数据的基础上（而不是依靠临床知识判断）使其得到改善。目前国外已经完成的 20 多个 PBE-CPI 研究表明，PBE-CPI 研究能够揭示在不同临床环境中、不同疾病状态下，多个诊疗组之间的疗效对比。

PBE-CPI 与 RCT 相比最大的优点就是能从实际医疗实践中发现那些能产生更好临床结局的干预措施且成本较低；同时，PBE-CPI 避免了观察性研究中仪器测量和无偏估计过程中技术因素的影响，提高研究结果的外部真实性。由于 PBE-CPI 主要关注那些有实际应用意义即可以用来提高医疗效果的结果，因此 PBE-CPI 实施中至关重要的就是必须使来自各科室的一线医生充分参与到研究中来，包括定义研究变量、执行研究计划、数据分析及把研究结果应用于临床实践。

表3 RCT 和 PBE-CPI 特点比较

项目	RCT	PBE-CPI
病人纳入/排除	排除可能使结果产生偏倚的病人，如患多种疾病者或出现并发症者，实际应用人群可能会被排除。仅 10% ~ 15% 的病人符合条件	用病情的严重程度来衡量多种疾病并发的情况和某疾病的严重性，一般不对研究对象做任何限制，所有病人都有资格入选，并对病人之间的差异进行估量，无病人被排除
医疗过程	依据干预方案，明确区分干预组和对照组医疗过程中的每一个重要因素，防止组间污染	对所有的处理和干预进行测量并详细记录，从现有的医疗记录中抽取信息
知情同意	必需	通常不需要知情同意*
临床结局	只针对主要结局：结局在证据的基础上变化	对多种结局进行评估：在数据分析的基础上得到的结局的改善
测量/记录	被测量和记录到的病人关键信息、医疗处理过程和临床结局信息有限	整体而全面
数据库	仅限于研究设计中规定的所需变量	全面而具体
结果	理论疗效：有条件限制的因果关系	实际疗效：关联，假定的因果关系
效度	强调内部真实性、外部真实性较低	强调外部真实性，获得外部真实性与内部真实性的最大平衡
研究假设	一种假设：在研究开始时即明确定义，范围较窄、较集中	多种假设：研究开始时广泛假设，在分析过程中可以不断提炼新的假设
地方性知识	不依赖于地方性知识	依赖于地方性知识；要求在职临床医生的参与
混杂因素	理论上与研究过程或结局无关—在研究设计中即已被排除	与纳入有关，而且确实会影响结局
可行性	昂贵、实际临床环境中难以广泛开展，且易受少部分依从性差者的影响	成本较低、适合在实际临床环境中广泛开展，不存在依从性的问题

注：* 在只需要从临床常规诊疗记录中获得数据和没有实验性干预措施的情况下不需要知情同意。

医疗服务提供者同时全程参与有利于互相学习交流，这对于把研究发现应用于临床实践，进而改善医疗服务过程非常有用。还要提及的一点是，创建 PBE-CPI 研究数据库是该研究方法实施的一个瓶颈，这需要建立在普遍的电子健康记录、有效的数据库链接、挖掘及使用工具、临床一线医生积极收集信息的基础之上。我国尚未形成以实践为基础的研究网络，因此在创建 PBE-CPI 研究数据库的初期需要花费大量的资金，其中包括对不同背景的一线医生进行“以实践为基础的证据”的概念培养和有关培训。当然，PBE-CPI 方法并不能代替 RCT，两种方法是互补关系。RCT 中证明有效的措施可以在 PBE-CPI 中得到进一步验证；而 PBE-CPI 又可作为新的 RCT 研究的起点。正如 Westfall 所说的：“以实践为基础的研究可以在研究和实践之间搭起一座桥梁，使研究所推荐的医疗服务应用到实践当中时能够真正地提高医疗水平。”

【参考文献】

[1] LeighAnne Olsen, Dara Aisner, and J. Michael McGinnis, eds. The Learning Healthcare System: Workshop Summary (IOM Roundtable on Evidence - Based Medicine), available at <http://www.nap.edu/catalog/11903.html>

[2] Brian L. Strom. Methodologic Challenges to Studying Patient Safety and Comparative Effectiveness. Med Care, 2007, 45 (Suppl 2): S13-S15

[3] 唐金陵, 江宇, 张宏伟. 随机对照试验. 见李立明主编《流行病学》第 6 版, 人民卫生出版社, 2007 年, 北京, 158 - 159

[4] Fayers PM, Jordhoy MS, Kaasa S. Cluster - randomized trials. Palliat Med, 2002, 26: 69-70

[5] Kathleen M. Mazor, James E. Sabin, Denise Boudreau, et al. Cluster Randomized Trials: Opportunities and Barriers Identified by Leaders of Eight Health Plans. Med Care, 2007, 45(Suppl 2): S29-S37

[6] Kathleen N. Lohr. Emerging Methods in Comparative Effectiveness and Safety: Symposium Overview and Summary. Med Care, 2007, 45 (Suppl 2): S5- CS8

[7] Susan D. Horn, Julie Gassaway. Practice - Based Evidence Study Design for Comparative Effectiveness Research. Med Care, 2007, 45 (Suppl 2): S50-S57