

比较效果研究 (CER) 方法在针灸临床研究中的应用

费宇彤 柴倩云 王聪聪 刘建平

(北京中医药大学循证医学中心,北京,100029)

摘要 文章简要介绍了比较效果研究(CER)的定义、常见研究方法及方法学要点,并较为系统地介绍了其在针灸领域的代表性研究。根据国际上的比较效果研究设计和评价的方法学工具,结合针灸临床实践和科研的特点,提出了针灸比较效果研究设计和报告的精确式与宏观式的方法学比较表单,以供使用者参考。

关键词 比较效果研究;CER;针灸;临床研究;方法学;疗效

Comparative effectiveness research methods for acupuncture

Fei Yutong, Chai Qianyun, Wang Congcong, Liu Jianping

(Center for Evidence - Based Medicine, Beijing University of Chinese Medicine, Beijing 100029, China)

Abstract We briefly introduced the concept, methods and key methodological considerations of comparative effectiveness research (CER), and provided a profile for important international acupuncture CERs and their contributions. Based on three world recognized evaluation scales and tools (the PRECIS, IOM 6 items, and the Efficacy - Effectiveness Continuum) for CERs, we developed a checklist for the design and appraisal of acupuncture CERs to illustrate the two types of acupuncture CERs to evaluate either a specific acupuncture intervention pack or a general concept of acupuncture practice in population base.

Key Words Comparative effectiveness research; CER; Acupuncture; Clinical research; Methodology; Clinical effectiveness

中图分类号:R2-03 文献标识码:A doi:10.3969/j.issn.1673-7202.2014.10.006

1 比较效果研究

以评价临床中实际应用的两种(或多种)干预措施综合临床效果差异为目的的研究被称作比较效果研究(Comparative Effectiveness Research, CER)^[1]。比较效果研究经常在“现实世界”中进行,评估比较两种或多种有可能成为临床最优治疗方案的不同医疗干预和策略的获益与风险。CER的目的是帮助医疗消费者、临床医生、政策制定者能够有科学根据地进行决策,以同时促进个体水平和群体水平的医疗卫生服务^[2]。CER方法在西方日益受到关注的主要原因来自西方国家缩减医疗费用负担的需要,和健康保险提供者、独立咨询机构以及临床医生需要解决临床实际问题的需要。

CER并非一种新的医学临床研究方法,而是包括一系列现代临床流行病学和循证医学领域的成熟方法。其共同特点是在临床医疗现实世界(real world)环境下,对有实际临床应用的不同干预(如药物、疗法等)进行相互比较,以期能够得到对临床有实际借鉴作用的研究结果。CER可以包括:1)非安慰剂对照的随机对照试验(Randomized Controlled Trials, RCTs),如实用型随机对照试验(Pragmatic RCT)、集团随机对照试验(cluster RCT)、交叉设计(cross-over design)等;2)

有对照的观察性研究(observational study),如队列研究(cohort study)、病例对照研究(case-control study)、历史对照的病例系列研究等;3)信息采集自疾病注册系统(disease registry)、电子病历数据库(electric medical record, EMR)、健康档案(electric health record, HER)的大样本观察性研究。

与解释性RCT(安慰剂对照)相比,CER更加注重:1)以患者为中心(patient-centeredness)的研究设计;2)关注患者的病情动态(patient dynamics);3)纳入存在临床异质性的患者(patient heterogeneity);4)关注个体化临床结局(patient values)。解释性RCT更加注重:1)以群组为中心(group-centeredness);2)关注起止点的结局数据(patient statics);3)纳入同质性患者(patient homogeneity)^[3]。

2 比较效果研究在针灸临床研究中的应用进展

比较效果研究是近10年国际针刺临床研究中的方法学热点,集中体现在如何判别取舍针刺临床研究的特异性疗效和综合性疗效方面。特异性疗效(efficacy)又翻译为“效力”,是指某一特定干预措施产生的具有特异性的治疗作用。特异性疗效通常通过解释性随机对照试验(与安慰剂对照)进行研究。综合性治疗

效果 (effectiveness) 是某一干预措施在临床现实世界中应用时观察到的治疗效果。这种效果可能受到各种混杂的影响,但能够较真实反映该干预应用于临床实践后的真实效果。综合性治疗效果通常可以通过比较效果研究来获得。

国际高质量大样本针刺临床试验最多的领域是疼痛的临床研究,进行了大量比较效果研究。

早期美国进行了两个大样本的针灸 CER。Vickers 等的研究(样本量 300 例)发现对于慢性头痛的治疗,针刺联合标准治疗组较之标准治疗组具有更好的效果^[4];同年, Berman 等发现针刺治疗膝关节炎效果明显优于健康教育和假针刺(样本量 580 例)^[5]。

此后,德国学者进行了逐项递进、互相承接的三个系列的大型针灸 CER,使针刺镇痛效果获得了国际认可,但同时也促使针刺安慰剂效应的凸显。在德国医疗保险系统的需求下, Claudia Witt 等著名学者牵头设计了三个系列的大型试验,其研究结果在 2006—2008 期间逐篇发表。第一个系列—ART (Acupuncture Randomized Trials),包括 4 个试验(每个试验约 300 例),对慢性腰痛^[6]、紧张性头痛^[7]、偏头痛^[8]、膝关节炎^[9]患者分别进行试验,比较针刺、假针刺和不治疗(等待名单)的效果差别,研究结果提示:对于慢性腰痛、紧张性头痛、偏头痛的治疗而言,针刺效果优于不治疗,但是与假针刺没有统计学意义;对于膝关节炎的治疗而言,针刺明显优于另外两个组。第二个系列—GERAC (German Acupuncture Trials),同样对上述四个病种进行研究,更加注重针刺治疗与临床常用治疗手段的疗效比较。GERAC 包括 4 个试验,其中 3 个试验为 3 组比较(每个试验约 1 000 例),比较针刺、假针刺及常规治疗(腰痛)^[10]或理疗(膝关节炎)^[11]或标准治疗(偏头痛)^[12];1 个 2 组试验(紧张性头痛,409 例)^[13],比较了针刺与假针刺。研究结果显示:对于膝关节炎、慢性腰痛,针刺治疗明显优于理疗或常规治疗,而与假针刺无差别;对于偏头痛,针刺治疗与标准治疗和假针刺均无差别;对于紧张性头痛,针刺与假针刺有统计学意义。第三个系列—ARC (Acupuncture in Routine Care Study),同样对上述 4 种疾病进行研究,属于典型的比较效果研究。ARC 包括 4 个试验,每个试验均为自由针刺联合常规治疗与常规治疗进行比较,其中 3 个试验^[14-16]样本量每个在 3 200~3 800 例之间,膝髌关节炎的试验^[17]为 712 例。研究结果均显示针刺与常规治疗联合使用明显优于常规治疗。

上述三个系列的大型德国针灸试验在国际上产生了巨大影响。其中某些研究样本量之大,涉及的针灸

医生之多均为历史之最。而且这些试验都是针对大约 4~5 种西方针灸科常见慢性疼痛而进行,又有共同的顶层设计,方法学质量相对较高,在世界范围内确证了针灸镇痛的临床疗效。但是由于这些试验同时揭示出真针刺与假针刺之间疗效差异不显著的现象,引发了一系列的方法学论证与创新。

此外,2013 年 9 月,著名的英国针灸临床研究者和方法学家 Hugh MacPherson 等在 Plos Medicine 发表了针刺治疗抑郁症的大型实用型随机对照试验^[18]。试验纳入了中重度抑郁症患者(BDI-II(Beck Depression Inventory-II) > 20 分)755 例,随机分为针刺+常规治疗组(302 例),咨询+常规治疗组(302 例)和常规治疗组(151 例)。其中,针刺组采用个体化治疗,由英国针灸专业学术组织内临床经验 3 年以上针灸师执行;咨询(counseling)由英国咨询与心理疗法协会会员执行,方法包括同情和倾听,并不给建议;常规治疗(usual care)由全科医生完成,在患者有诉求时给予治疗。针刺治疗大约每周 1 次,共 12 周。研究结果显示针刺组效果与咨询组没有显著性差异,但是均明显优于常规治疗组。此项研究是近年国际学者进行的很有影响力的非痛症领域的大型临床研究。

国内学者在针灸的比较效果研究中也做了大量工作,涉及到不同针灸疗法之间的比较,针灸与中药或西医治疗手段之间的比较,针灸联合中药与中药或西医治疗手段的比较等。相比之下,国内的临床研究不少,可能也不乏高质量研究,但是在国外高影响因子学术期刊发表的大型 CER 还非常有限,影响力也没有达到应有的高度。随着国内针灸临床研究学术水平的不断提高,近年可能会有更多高质量 CER 得到国际同行的认可。

3 针灸 CER 设计和报告中的方法学要点

根据 PRECIS (Pragmatic - explanatory Continuum Indicator Summary)^[19]、IOM (Institute of Medicine) 6 条标准^[20]和改良的效力-效果连续体 (Efficacy - Effectiveness Continuum)^[21],我们提出针灸 CER 方法学比较表单(表 1),以供针灸 CER 设计和报告时参照使用。

总之,非药物疗法的临床研究较之药物疗法而言,干预措施的操作性和技术性是整个研究设计、执行和结果推广的关键点。其中,实施者技术的规范性、稳定性、一致性,实施者经验、技能的要求必须加以重视。对于比较效果研究亦是如此。特别需要注意的是,患者异质性强的 CER 在设计时需要更大样本量,CER 设计中患者的分层标准对亚组分析至关重要。

表 1 针灸 CER 方法学比较

方法学特征	精确式 CER	宏观式 CER
适用性	具体详细的针灸治疗方案与另外一种临床潜在最佳治疗比较,更有利于针灸技术的传承与推广	笼统的针灸治疗方案,甚至不涉及具体针灸方案,与另外一种临床潜在最佳治疗比较。更有利于探索针灸的普遍疗效
患者	纳入范围狭窄;使用逐步递进的患者筛选方法 1. 仅纳入不良结局最高发的人群; 2. 进一步限制于预计对试验措施高度敏感的人群; 3. 进一步筛选预计具有较好依从性的患者(如既往就诊履约良好等)	纳入范围宽泛: 将所有具有目标疾病(健康状况)的人群全部纳入,不限制其预期风险、反应度、共患病或者既往依从性
针灸干预措施	尽量标准化 1. 无变化的试验干预措施,对各个要素(腧穴选择方法、入针法、行针法、补泻手法、留针法、出针法、腧穴针刺顺序、时间)都有严格限制; 2. 针灸医生教育背景、临床工作经验及针刺操作习惯接近(有统一入选标准); 3. 标准化针灸操作的培训及严格量化考评以保证针灸操作能满足方案要求; 4. 规范化针刺过程中的医患交流方式、内容、时间,以减少过度热情或冷淡的医患交流; 5. 尽量减少临床试验中针灸师的更迭	尽量日常化 1. 对于如何使用试验干预措施的指导非常灵活,实施者有相当大的自由决定处方和治疗; 2. 在符合当地卫生法规的前提下,对参加试验的针灸医生的教育背景、临床工作经验及针刺操作习惯等不做特殊要求; 3. 对针灸操作的要求可仅通过研究方案传达,可以没有硬性培训
联合干预措施(与针灸联使用)	如果有,则尽量标准化	如果有,则尽量标准化
对照干预措施	尽量标准化的临床潜在最佳治疗方案(措施) 尽量客观化	尽量实用化的临床潜在最佳治疗方案(措施) 以患者为中心
结局指标	1. 结局是干预措施的已知、直接、立效结局; 2. 主要结局是可以客观测量的,对患者而言有临床意义的; 3. 结局的测量在日常临床中可以实现(无需专门检测或培训)	1. 主要结局是可以客观测量的,对患者而言有临床意义的; 2. 结局的测量在日常临床中可以实现(无需专门检测或培训)
研究背景	试验条件 1. 特异性很强的医疗/试验机构 2. 日常诊疗环境	日常条件 1. 日常诊疗环境

CER 终极目的是在现实临床情境和现实患者群中,比较不同医疗管理策略的潜在益处与害处。因此,CER 的报告要强调:第一,必须及时满足当前医疗决策者的需求;第二,必须客观、科学、严谨,严格控制由于个人和商业因素所带来的偏倚;第三,实施过程必须公开、透明。而且,为了说明 CER 研究的外推性,研究报告中应该包括如下内容:临床研究环境(条件)与临床实践环境(条件)的差异性说明和讨论;患者依从性的说明和讨论;以及医生执行研究方案的依从性的说明和讨论。

4 中医(针灸)领域 CER 研究的展望

中医(针灸)流派的疗效对比评价研究,个体化辨证论治的动态评价研究(随证调方难以制作安慰剂),中医治则治法的疗效评价研究,名老中医经验传承,针灸临床疗效评价研究(安慰针、安慰灸难以实现),其他中医疗法,如推拿、拔罐、食疗、医学气功(如太极)、养生等的临床疗效评价均可以酌情优先考虑使用 CER。在选用 CER 进行研究设计时,应该根据研究目的选择性使用精确式 CER 或者宏观式 CER。

参考文献

- [1] Nicholas F. Marko, Robert J. Weil. An Introduction to Comparative Effectiveness Research[J]. Neurosurgery, 2012, 70(2): 425-34.
- [2] Sox HC, Greenfield S. Comparative effectiveness research: a report from the Institute of Medicine[J]. Ann Intern Med, 2009, 151(3): 203-205.
- [3] Mikel Aickin. Comparative Effectiveness Research and CAM[J]. The Journal of Alternative and Complementary Medicine, 2010, 16(1): 1-2.
- [4] Vickers AJ, Rees RW, Zollman CE, McCarney R, Smith CM, Ellis N, Fisher P, Van Haselen R. Acupuncture for chronic headache in primary care: large, pragmatic, randomised trial[J]. Bmj, 2004, 328(7442): 744.
- [5] Berman BM, Lao L, Langenberg P, Lee WL, Gilpin AM, Hochberg MC. Effectiveness of acupuncture as adjunctive therapy in osteoarthritis of the knee: a randomized, controlled trial[J]. Ann Intern Med, 2004, 141(12): 901-910.
- [6] Brinkhaus B, Witt CM, Jena S, Linde K, Streng A, Wagenpfeil S, Irnich D, Walther HU, Melchart D, Willich SN. Acupuncture in patients with chronic low back pain: a randomized controlled trial[J]. Arch Intern Med, 2006, 166(4): 450-457.
- [7] Melchart D, Streng A, Hoppe A, Brinkhaus B, Witt C, Wagenpfeil S, Pfaffenrath V, Hammes M, Hummelsberger J, Irnich D, et al. Acupuncture in patients with tension-type headache: randomised controlled trial[J].

- Bmj,2005,331(7513):376-382.
- [8] Linde K, Streng A, Jurgens S, Hoppe A, Brinkhaus B, Witt C, Wagenpfeil S, Pfaffenrath V, Hammes MG, Weidenhammer W, et al: Acupuncture for patients with migraine: a randomized controlled trial [J]. *Jama*, 2005, 293(17):2118-2125.
- [9] Witt C, Brinkhaus B, Jena S, Linde K, Streng A, Wagenpfeil S, Hummelsberger J, Walther HU, Melchart D, Willich SN: Acupuncture in patients with osteoarthritis of the knee: a randomised trial [J]. *Lancet*, 2005, 366(9480):136-143.
- [10] Haake M, Muller HH, Schade - Brittinger C, Basler HD, Schafer H, Maier C, Endres HG, Trampisch HJ, Molsberger A: German Acupuncture Trials (GERAC) for chronic low back pain: randomized, multicenter, blinded, parallel - group trial with 3 groups [J]. *Arch Intern Med*, 2007, 167(17):1892-1898.
- [11] Scharf HP, Mansmann U, Streitberger K, Witte S, Kramer J, Maier C, Trampisch HJ, Victor N: Acupuncture and knee osteoarthritis: a three-armed randomized trial [J]. *Ann Intern Med*, 2006, 145(1):12-20.
- [12] Diener HC, Kronfeld K, Boewing G, Lungenhausen M, Maier C, Molsberger A, Tegenthoff M, Trampisch HJ, Zenz M, Meinert R: Efficacy of acupuncture for the prophylaxis of migraine: a multicentre randomised controlled clinical trial [J]. *Lancet Neurol*, 2006, 5(4):310-316.
- [13] Endres HG, Bowing G, Diener HC, Lange S, Maier C, Molsberger A, Zenz M, Vickers AJ, Tegenthoff M: Acupuncture for tension - type headache: a multicentre, sham - controlled, patient - and observer - blinded, randomised trial [J]. *J Headache Pain*, 2007, 8(5):306-314.
- [14] Witt CM, Jena S, Selim D, Brinkhaus B, Reinhold T, Wruck K, Liecker B, Linde K, Wegscheider K, Willich SN: Pragmatic Randomized Trial Evaluating the Clinical and Economic Effectiveness of Acupuncture for Chronic Low Back Pain [J]. *Am J Epidemiol*, 2006, 164(5):487-496.
- [15] Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN: Acupuncture for patients with chronic neck pain [J]. *Pain*, 2006, 125(1-2):98-106.
- [16] Jena S, Witt CM, Brinkhaus B, Wegscheider K, Willich SN: Acupuncture in patients with headache [J]. *Cephalalgia*, 2008, 28(9):969-979.
- [17] Witt CM, Jena S, Brinkhaus B, Liecker B, Wegscheider K, Willich SN: Acupuncture in patients with osteoarthritis of the knee or hip: a randomized, controlled trial with an additional nonrandomized arm [J]. *Arthritis Rheum*, 2006, 54(11):3485-3493.
- [18] MacPherson HI, Richmond S, Bland M, Brealey S, Gabe R, Hopton A, Keding A, Lansdown H, Perren S, Sculpher M, Spackman E, Torgerson D, Watt I: Acupuncture and counselling for depression in primary care: a randomised controlled trial [J]. *PLoS Med*, 2013, 10(9):e1001518. doi:10.1371/journal.pmed.1001518.
- [19] Daniel L. Riddle. et al. The Pragmatic - Explanatory Continuum Indicator Summary (PRECIS) instrument was useful for refining a randomized trial design: Experiences from an investigative team [J]. *Journal of Clinical Epidemiology*, 2010, 63(11):1271-5.
- [20] Harold C. Sox. Initial National Priorities for Comparative Effectiveness Research; Report Brief. 2009; <http://www.iom.edu/Reports/2009/ComparativeEffectivenessResearchPriorities.aspx>.
- [21] Witt CM, Michalsen A, Roll S, Morandi A, Gupta S, Rosenberg M, Kroppass L, Stapelfeldt E, Hissar S, Müller M, Kessler C. Comparative effectiveness of a complex Ayurvedic treatment and conventional standard care in osteoarthritis of the knee - study protocol for a randomized controlled trial [J]. *Trials*, 2013, 23(14):149. doi:10.1186/1745-6215-14-149.

(2014-09-21 收稿 责任编辑:洪志强)

(上接第 1273 页)

在此前提下,从倾向指数获得的干预效果的估计才是无偏的。而且,暴露组和对照组的组间特征变量要有足够的重叠,即有相当大的共同支持域(common support region)。鉴于特征变量的混杂效应广泛的存在于非随机的观察性研究中,倾向指数的应用需要有大样本量作为前提,从而保证暴露组和对照组之间倾向指数分布的均衡性。样本量越大,特征变量之间的均衡性越好。

总之,在中医药治疗的临床评价研究中,非随机的大样本观察性研究数据对于治疗效果的因果推论仍然具有很大的价值和潜力。倾向指数的应用将会更好的控制可观测的特征变量的差异,获得最佳效应的估计值。

参考文献

- [1] 刘建平. 传统医学证据体的构成及证据分级的建议 [J]. *中国中西医结合杂志*, 2007, 27(12):1061-1065.
- [2] Shadish, W. R., Cook, T. D., & Campbell, D. T. *Experimental and quasi-experimental designs for generalized causal inference* [M]. Boston: Houghton Mifflin, 2002.
- [3] Neyman, J. . Statistical problems in agricultural experiments [J]. *Supplement to the Journal of the Royal Statistical Society*, 1935, 2(2):107-180.
- [4] Rubin, D. B. . Estimating causal effects of treatments in randomized and nonrandomized studies [J]. *Journal of Educational Psychology*, 1974, 66(5):688-701.
- [5] 陈蔚文. *中药学* [M]. 2 版. 北京: 人民卫生出版社, 2013.
- [6] 费宇彤, 刘建平. 国际草药 CONSORT 声明及中药临床试验报告规范化问题思考 [J]. *中国中药杂志*, 2008, 33(1):89-94.
- [7] 刘建平. 中医药临床试验的方法学问题与挑战: 循证医学的观点 [J]. *中西医结合学报*, 2006, 4(1):1-6.
- [8] 沈玉红, 彭真, 张珂良. 中药临床试验现状及对策分析 [J]. *中国新药杂志*, 2013, 22(20):2365-2368.
- [9] 范昕, 李康, 马葆华, 等. 中药临床试验中若干医学伦理学问题 [J]. *中医药学报*, 2012, 40(1):7-10.
- [10] Rosenbaum P R. Rubin D B. The central role of the propensity score in observational studies for causal effects [J]. *Biometrika*, 1983, 70(1):41-55.
- [11] 辛涛, 姜宇. 教育科学的因果推论: 困境与超越 [J]. *清华大学教育研究*, 2013, 34(3):1-6.

(2014-09-21 收稿 责任编辑:洪志强)