

·方法学与临床评价·

随机对照试验设计对针灸临床研究设计的启示

吴曦¹ 梁繁荣¹ 李瑛¹ 郑晖¹ 任玉兰¹ 蒋可²

(1. 成都中医药大学针灸推拿学院, 四川省成都市十二桥路37号, 610075; 2. 中国中医科学院西苑医院针灸科)

[摘要] 循证医学的兴起对临床研究的质量提出了高水平的要求。随机对照试验被循证医学认为是证据级别和论证强度很高的临床研究。列举了若干种重要的、具有启发性的随机对照试验设计类型,并展望了针灸临床研究设计发展的方向,期望能为针灸临床研究设计提供一定程度的借鉴。

[关键词] 随机对照试验; 针灸; 临床研究设计

Enlightenment of Randomized Controlled Trial on the Design of Clinical Research of Acupuncture and Moxibustion

WU Xi¹, LIANG Fanrong¹, LI Ying¹, ZHENG Hui¹, REN Yulan¹, JIANG Ke²

(1. College of Acupuncture and Massage, Chengdu University of TCM, Chengdu 610075; 2. Xiyuan Hospital, China Academy of Chinese Medical Sciences)

ABSTRACT The rising of evidence-based medicine requires high quality for clinical research and randomized controlled trial is considered as clinical research of the most level of evidence degree and confidence activity. This article illustrates several important and instructive designs of randomized controlled trials, which can give some reference for clinical research on acupuncture. And the developing direction of future clinical research on acupuncture is discussed.

Key Words Randomized controlled trial; Acupuncture and moxibustion; Design of clinical research

循证医学被誉为21世纪的临床医学,其兴起对临床研究的质量和水平提出了更高的要求。对各种类型的临床研究,循证医学最关注研究结果的真实性和可靠性。随机对照试验(RCT)被认为是证据级别和论证强度很高的试验类型。下面以Alejandro R. Jadad教授所著第1、2版 *Randomized Controlled Trials* 为基础^[1-2],并参考其他针灸临床研究设计介绍^[3],列举说明对针灸临床研究设计具有启发性的RCT设计,以供针灸临床研究设计者参考借鉴。

1 解释性和实效性试验

解释性试验(explanatory trials)是评价一项干预措施是否有效,以及如何起效的研究。经典的解释性试验需要设立严格的纳入标准,长时间的洗脱期,设立安慰剂对照组,采用固定的干预方法,意向性分析,关注硬性指标,因此是倾向于对干预措施的内部真实性进行评价。

实效性试验(pragmatic trials)又称为管理试验(management trials),是评价在模拟临床实践的环境下干预措施是否有效,以及所有的无论好坏的应用后果。实效性试验往往设立更加宽松的纳入标准,如同临床

医师每日在临床实践中遇到的患者那样,采用更加灵活的对照(不一定是安慰剂,可能是其他有效的干预措施),灵活的干预方法(比如第1周每天针刺1次,从第2周开始隔日针刺1次),采用符合方案集分析,在结局指标上可能是生活质量等软指标。该设计更倾向评价干预措施的外部真实性。需要说明的是,解释性试验和实效性试验是合理并互为补充的,大部分RCT是对两种设计的联合运用。

2 效力和效果试验

效力试验(efficacy trials)指在理想的、实施严格控制的条件下,给予治疗措施最优实施的情况下,判断治疗措施是否有效的临床试验。其与解释性试验类似,在很多情况下两者被等同,该设计的特点在于研究者不太关注干预措施是如何起效的,其主要目标是尽量保证受试者对干预方案的良好依从性,以达到对干预措施在受试对象中是否有效的评价。

效果试验(effectiveness trials)指在与现实临床实践环境相似的情况下,判断治疗措施是否有效的临床试验。经典的效果试验是评价已经证实了效力的干预措施,在日常临床环境下给予异质性人群的疗效,因此是倾向于评价干预措施的外部真实性。其纳入标准宽松,干预方法灵活,并允许受试者接受或拒绝研究者提供的干预措施。例如表1中的针刺随机对照研究^[4]说

明,研究设计运用的关键在于根据回答的临床问题,设计各个要素的最佳组合。

表1 针刺治疗持续的非特异性腰痛随机对照研究设计	
项目	设计方案
研究目标	短期疗程的传统针灸在初级医疗保健中能否改善持续的非特异性低背痛患者的长期结局
研究设计	实效性研究,开放的 RCT 试验,采用计算机区组随机,评价时采用盲法
实施环境	英国约克郡3个私人诊所和18个初级医疗保健门诊
受试者	241个18~65岁的成年患者,经过初级医疗保健医师评估,排除可能患有脊髓疾病、腰椎间盘突出等特异性腰痛和正在接受针灸治疗,证实患有非特异低背痛,病程4~52周
干预措施	160例针刺组由6名临床实践至少3年以上的英国针灸协会的注册针灸师,在3个月内对患者进行10次个体化针刺辨证治疗,80例对照组由初级医疗保健医师给予英国国家医疗服务系统(NHS)标准治疗
结局指标	主要结局指标:第12和24个月时简明健康调查量表(SF-36)中躯体疼痛维度,次要结局指标:Oswestry 疼痛残疾指数、McGill 当前疼痛指数,SF-36 总体健康维度
研究结果	213例完成12个月随访,与对照组相比针刺的疗效是5.6分(95%CI: -0.2~11.4),182例完成24个月随访,疗效为8.0分(2.8~13.2);Oswestry 疼痛残疾指数、McGill 当前疼痛指数有差异,但无统计学意义;总体健康维度无差异
安全性评价	无患者报道的严重不良事件,23%(30/133)患者报道了治疗期间严重干扰生活的短暂低背痛加重,86%患者将针刺治疗作为放松,91%患者欢迎再次接受针刺
结论	低强度证据显示12个月时针刺对非特异性低背痛的疗效,更强的证据显示24个月时针刺的疗效;熟练传统针灸医师的短期治疗是安全的,并为低背痛患者所接受

3 优效性、非劣效性和等效性试验

优效性试验(superiority trials)指主要研究目为显示所研究的干预措施疗效优于对照措施(安慰剂或阳性对照)的试验。非劣效性试验(noninferiority trials)指主要研究目的为显示干预措施的疗效在临床意义上不差于(非劣于)对照措施的试验。等效性试验(equivalence trials)指主要研究目的是显示两种或多种干预措施疗效差异的大小在临床上并无重要性的试验,通常通过显示真正的差异位于临床上可以接受的等效的界值之间来证实。后面两种设计随着临床医学的不断发展而得到越来越多的应用。等效性试验设计举例如表2^[1]。

4 平行和交叉试验

平行试验(parallel trials)是指每组的受试者只接受研究的一种干预措施。大部分RCT采用这种设计。交叉试验(crossover trials)是指每组的受试者在连续的阶段内给予研究的所有干预措施,接受每种干预措施的顺序是随机决定的。因此,平行试验是患者之间疗效的比较,而交叉试验是每个患者的自身前后对照。交叉试验以每例受试者自身为对照,可消除个体的差异,减少样本含量,每例受试者均有接受两种干预措施的机会,

减少了意愿偏倚。近年来受到广泛重视的单个患者试验(single-patient trials)就是一种特殊的交叉试验,其经典设计如表3^[6]。据研究,经典的4次交叉试验具有80%的效能检验出20%的差异,且节约时间和成本。美国医学会在给临床医生的治疗决策推荐中,证据强度最高的就是单个患者随机对照试验,认为其证据的强度高于随机对照试验的系统评价^[7]。

表2 针刺预防偏头痛发作的多中心随机对照临床试验设计	
项目	设计方案
研究目标	比较半标准化针刺、半标准化假针刺以及标准预防药物对偏头痛发作后26周降低患者头痛天数的疗效
研究设计	前瞻性、随机、多中心、双盲(盲患者和观察者)、平行、对照临床试验
假设检验 H ₀	3个治疗组可能取得的疗效无差异(界值=2天)
受试者	960例每月发作2~6次的偏头痛患者(有详细的纳入标准和排除标准)
干预措施	针刺组(313例)和假针刺组(339)由至少完成140h培训和2年临床工作经验的针灸医师按照传统的针灸理论,在6周内给予10次治疗。药物组(308例)按照德国偏头痛和头痛学会指南给予持续的药物治疗性治疗(β受体阻滞剂、西比灵或丙戊酸)
结局指标	主要结局指标:头痛日记反映的基线期的平均头痛天数与随机后23~26周平均头痛天数的差值。次要结局指标:有效率(定义为同基线相比头痛天数下降50%及以上)等
统计分析	采用 SAS 9.1 软件,主要结局指标进行意向性分析(ITT)和符合方案集分析(PP),呈非正态分布采用 Kruskal-Wallis 检验。对次要结局指标的探索性分析采用 ITT 分析,根据定性或定量变量分别采用 χ ² 检验或 Kruskal-Wallis 检验
研究结果	794例完成 ITT 分析,443例完成 PP 分析,针刺组疗效 2.3 天(95%CI: 1.9~2.7 天),假针刺组 1.5 天(1.1~2.0),药物组 2.1 天(1.5~2.7),3 组与基线相比差异有统计学意义(P<0.0001),但组间无差异(P=0.09)。3 组有效率分别为 47%、39%、40%,组间比较无差异(P=0.133)
安全性评价	与针刺有因果关系的严重不良事件 1 起,3 组的不良事件发生比例分别为 70(22%)、75(23%)、59(31%),不良事件主要是神经系统、骨骼肌以及结缔组织问题
结论	研究结果显示,对偏头痛的预防,半标准化针刺、半标准化假针刺和标准预防药物的疗效无差异

表3 单个患者试验设计方案		
阶段	时间	干预方法
第一交叉阶段	1~4天	A
	5~8天	B
第二交叉阶段	9~12天	B
	13~16天	A
第三交叉阶段	17~20天	B
	21~24天	A
第四交叉阶段	25~28天	A
	29~32天	B

交叉试验并不适用于所有的临床情况,其实施有一些必要的条件:研究对象必须是慢性的、不可治愈的疾病;干预措施的疗效起效快,持续时间短;每次干预之前疾病或受试者的状态要求稳定,以保证可比性。

5 参与者选择试验

参与者选择试验 (participant preferences) 包括 Zelen's 设计、Wennberg's 设计、Brewin-Bradley 设计以及 Comprehensive cohort 设计等, 随着生物医学模式的变化, 这些设计越来越受到重视。

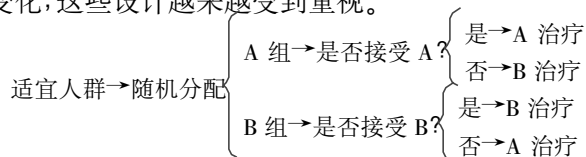


图1 Zelen's 设计

Zelen's 设计的特点在于患者在知情同意之前就被随机分配到了标准治疗组或实验干预组, 现在运用较多的模式是完全随机知情同意设计 (double randomised consent design), 如图1所示。该设计的优点在于所有的适宜人群都被纳入了试验, 而且保证了对受试者真实疗效的评估。主要的缺点是只能进行开放性研究, 在统计的时候无法进行 ITT 分析。

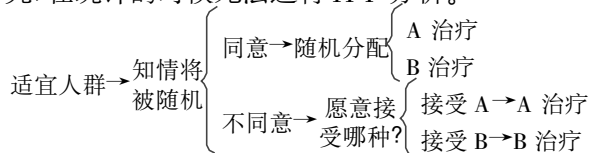


图2 Comprehensive cohort 设计

Comprehensive cohort 设计很适合受试者或研究者对某项干预措施有很强烈的愿望而导致高比例的适宜人群拒绝随机化的情况, 这时研究实际上就是前瞻性序列研究加上低比例参与者的 RCT 研究。这种研究的主要局限在于序列研究和 RCT 研究结果的差异总可以用两者基线的不一致进行解释。

6 展望

21世纪以前, 针灸研究主要以叙述性研究——系列病例分析、个案报道、专家经验、专家评述、杂志评论等为主, 反映了临床经验积累的水平和对相关领域研究结果的关注程度, 但由于多为回顾性研究, 缺乏严格的设计, 其论证强度不高。在现代临床流行病学建立以前, 临床试验传统采用的是非随机同期对照试验 (non-randomized concurrent control study), 即由医生或患者人为地决定纳入试验组或对照组的同期对照试验。至今针灸临床研究中仍有相当部分非随机同期对照试验, 由于受试对象的基线状态不明, 以及未实施盲法评价, 造成已知和未知的偏倚对研究结果真实性的影响。因此, 遵循随机化、对照和重复的基本原则进行临床研究设计, 避免严重的偏倚, 是今后针灸临床研究设计的发展方向。

21世纪是融合医学发展的重要时期, 在针灸临床疗效和作用机制引起全球医学界广泛关注和争论的背景下, 针灸临床研究的设计是影响其疗效评价结果的

重要因素。本文介绍了几种经典的、具有启发性的随机对照研究设计的方法, 它们各具特点, 可供今后针灸临床研究设计作为参考。例如, 现在被证实论证强度很高的单个患者试验, 非常符合针灸个体化治疗、检测和评估的特点, 可实现整体疗效评价和个体疗效评价的有机结合, 非常值得探索。再如针灸与药物比较的临床研究, 考虑到针灸具有毒副作用小, 医疗成本低的优点, 在疗效评价时就可以采用等效性或非劣效性设计。此外, 随着以患者为中心的医学实践模式的出现, 参与者选择试验的结果也具有相当的价值和意义。但是需要说明的是, 随机对照的研究设计并不适用于所有的研究目标, 如针对针刺不良反应的研究。在经济条件有限、患者依从性很低或退出率很高、干预至结果测量的时间太长等情况下, 其他的研究设计如病例对照研究、队列研究会更加合适。

在今后的针灸临床研究中, 我们可首先开展高质量的观察性研究。在此基础上, 对有证据的干预方法或内容进行小范围的试验性研究。在获得有益结果后, 再开展更加“昂贵的”大样本随机对照试验。在设计上注意突出针灸辨证论治和个体化治疗的特点, 从针灸的效力、效果、成本、成本-效果等多层次、多角度、系统地阐释针灸疗效的科学性、安全性和推广应用的优劣势。

参考文献

- [1] Jadad AR (Ed). Randomized controlled trials [M]. London: BMJ Publishing Group, 1998: 10-27.
- [2] Jadad AR, Enkin MW (Eds). Randomized controlled trials: questions, answers and musings, 2nd edition [M]. Chichester: John Wiley and Sons Ltd, 2007: 12-28.
- [3] Witt CM. Efficacy, effectiveness, pragmatic trials - guidance on terminology and the advantages of pragmatic trials [J]. Forsch Komplementmed, 2009, 16(5): 292-294.
- [4] Thomas KJ, MacPherson H, Thorpe L, et al. Randomised controlled trial of a short course of traditional acupuncture compared with usual care for persistent non-specific low back pain [J]. BMJ, 2006, 333(7569): 623.
- [5] Diener HC, Kronfeld K, Boewing G, et al. Efficacy of acupuncture for the prophylaxis of migraine: a multicentre randomised controlled clinical trial [J]. Lancet Neurol, 2006, 5(4): 310-316.
- [6] Reitberg DP, Del Rio E, Weiss SL, et al. Single-patient drug trial methodology for allergic rhinitis [J]. Ann Pharmacother, 2002, 36(9): 1366-1374.
- [7] Guyatt GH, Haynes RB, Jaeschke RZ, et al. Users' Guides to the Medical Literature. Evidence-based medicine: principles for applying the Users' Guides to patient care. Evidence-Based Medicine Working Group [J]. JAMA, 2000, 284(10): 1290-1296.

(收稿日期: 2010-04-20; 修回日期: 2010-05-26)

[编辑: 崔京艳]