

· 书刊评介 ·

中医药学术期刊随机对照临床试验文章评阅及建议*

詹思延¹ 唐金陵² 谢立亚² 吴涛¹ 贝东¹ 李立明¹

随着循证医学(evidence-based medicine, EBM)的出现,关于临床疗效的科学证据日益受到重视。众所周知,随机对照试验(randomized control trial, RCT)是评价医疗措施疗效最严谨的方法⁽¹⁾。20 年来我国在中医药疗效评价方面做了许多工作,但由于中西医体系不同,临床试验在中医中的应用亦有其特殊性,有关经验亟待总结。本研究评阅了有代表性的中医药学术期刊 28 种,旨在估计所发表的 RCT 总数, RCT 的上升趋势以及主要评价的疾病,为系统地收集、整理和传播中医临床疗效的 RCT 证据打下基础,即估计工作量和研究的重点⁽²⁾。同时详细评阅了中国中西医结合杂志自创刊到 1996 年发表的中医 RCT 文章,着重分析了几个主要影响 RCT 疗效估计的准确性和精确性,以及影响读者正确解释和利用 RCT 结果的因素,从而为今后中医 RCT 评价提出改进建议。

资料与方法

1 研究对象及定义 分层随机选择综合性国家级中医药学术期刊 4 种、省级和中医学院(大学)学报类中医药学术期刊各 10 种,专业杂志(如针灸科)4 种,共计 28 种,从 1996 年 12 月开始回溯 20 年,从中选取所有中医药疗效评价的文章,逐一评阅。RCT 文章入选标准:(1)以患者作为研究对象,并随机分组;(2)使用中医作为治疗手段(包括中西医结合的治疗方法);(3)有一个或几个适当的对照组;(4)以观察疗效为主要目的。

2 评阅方法和内容 由经过培训的调查员两人一组双检索、收集文章。参阅 EBM 工作组对 RCT 质量的评阅方法^(3,4),制定规范评阅表,摘录有关疾病、治疗方法及与 RCT 质量相关的指标等。质量评价的指标包括:随机分组的描述,盲法和安慰剂的使用,试验组与对照组可比性,不依从、失访和更换干预措施的描述,统计学检验和统计量的使用,副作用说明,结果评估的客观性等。评阅过程中,由主要作者随机抽查 23

篇(5.5%)共 989 项内容,与调查员比较项目不一致率为 1.9%。

3 统计分析方法 用 Epi-Info 软件建立数据库,SPSS 软件进行数据分析。

结 果

1 RCT 数量、增长趋势及研究的主要疾病

1.1 RCT 的数量 28 种杂志在 1997 年前的 20 年中,共发表 2938 篇 RCT 文章,包括摘要和短篇报道。其中,全国性杂志 4 种,发表 RCT 1328 篇(45.2%);省级杂志 10 种,710 篇(24.2%);学报类杂志 10 种,385 篇(13.1%);针灸杂志 4 种,515 篇(17.5%)。依此估计全国 100~120 种中医药学术期刊发表 RCT 总数在 7500 篇(95%可信限为 6000~9000)。进一步与中医文献分析摘录系统(TCMLARS,中医研究院情报所)中查到的 RCT 进行比较和推算,西医期刊发表的中医药 RCT 相当于中医药学术期刊发表的 RCT 总数的 20%。因此,1997 年前中国医学期刊共发表近 10 000 篇评价中医药疗效的 RCT 文章。

1.2 RCT 的增长趋势 1979 年以前,无 RCT 报道,至 1984 年底 28 种杂志共发表 90 篇。其后累积 RCT 总数每 2~3 年翻一番。

1.3 RCT 所研究的主要疾病 24 种综合杂志 2423 个 RCT 中,研究最多的 10 种疾病依次为冠心病、脑卒中、病毒性肝炎、胃溃疡、婴幼儿腹泻、高脂血症、高血压、上消化道出血、糖尿病、肺炎。10 种合计占 2423 个 RCT 的 23%。

2 中国中西医结合杂志发表的全文型 RCT 中国中西医结合杂志发表的 RCT 数量是其他杂志平均发表数量的 10 倍左右,且 RCT 质量也优于其他杂志⁽⁵⁾,因此对该杂志全文型 RCT 的特点、质量及发表偏性进行详细分析。

2.1 RCT 占有所有有关临床疗效文章的比例 中国中西医结合杂志自创刊起至 1996 年第 12 期止,共发表全文型有关临床疗效的文章 1687 篇,其中符合本研究入选标准的全文型 RCT 414 篇,占 24.5%。(另有 455 篇 RCT 为短篇报道,占总 RCT 的 52.3%)。随着年代的增长,全文型 RCT 占临床疗效文章的比重

* 本课题由香港中文大学资助(No. 2040518)

1. 北京医科大学流行病学教研室(北京 100083);2. 香港中文大学社区与家庭医学系

越来越大,1981~1985、1986~1990 和 1991~1996 年 3 个时期分别为 8%,19% 和 45%。鉴于 RCT 是评价临床疗效最严谨可靠的方法,RCT 在相关文章中的比重有待增加。此类文章应尽可能以全文形式报道,因为短篇报道使读者难以判断方法的科学性和结果的可用性。

2.2 治疗方案的组成 RCT 文章中所使用的治疗措施以药物为主,占 90%,而各种药物配方中 81% 为单方,即统一剂型和剂量的治疗措施,如中成药。

2.3 影响 RCT 质量因素的分析 见表 1。

2.3.1 随机方法描述 正确使用随机分组是取得比较组间初始可比性的保证。83% 的文章未描述随机分组方法,操作是否得当难以判断。

2.3.2 对照的选择 90% 以上的 RCT 以疗效明确的西医和(或)中医作为对照,是合理的。如果与疗效不明的治疗方案比较,得不出任何有价值的信息,应当彻底杜绝。如果没有疗效肯定的对照组,应使用空白对照或安慰剂对照。

2.3.3 组间可比性分析 59% 的文章没有提及基线一致性比较,虽然近年来进行这种分析的比例有所升高。评价一个治疗方案时,治疗组和对照组之间除治疗方案外,任何其他能影响疾病转归的因素,如年龄、性别、病情,应严格可比。随机分组就是为了达到这个目的。但当样本较小时,两组的可比性仍可能不好。因此在 RCT 样本较小时,检查两组基线可比性是极其必要的。若两组间有明显差异,在结果分析时应控制混杂。

2.3.4 盲法的应用 盲法旨在减少或避免由于患者、治疗者和(或)观察者由于知道分组情况而在试验过程中可能带来的各种偏性。研究发现,盲法在历年来 RCT 文章中的应用比例都很低(14.7%)。尤其是双盲法的应用,414 篇文章中仅有 28 篇提及(6.7%),且 27 篇以单方作为干预措施;使用辨证施治干预措施的 RCT 应用双盲法虽然有困难,但使用单盲(盲患者或结果观察者)是可行的,但 76 篇此类 RCT 中仅有 6 篇(7.9%)提及使用了单盲法。

2.3.5 样本量 对于疗效中等及以下的治疗方法的评价,样本量足够大是得出精确结论的重要因素之一。鉴于中医多用于慢性病治疗,且大多数只有中等疗效,因此样本一般要求较大^[6]。本次评阅文章中的样本量多在 50~250(66.2%),样本大于 300 的研究只有 40 篇,占 9.5%。另外有 9 篇文章(2.2%)的样本量小于 20。

2.3.6 不依从、失访和更换干预措施 RCT 中

这些事件或多或少是不可避免的,尤其未用盲法时,它们势必破坏组间可比性。因此研究者应明确报告它们发生的多少,并在结果分析时采用相应的措施,如目前常用的实效分析^[7]方法以保证随机分组不被破坏。但 RCT 文章中提及此类问题的比例很低。

2.3.7 统计分析 研究结果有统计学检验的文章比例达 90%,但提及采用何种统计量的文章比例较低。这会导致可能由于错用统计量,且无法被编审及读者发现,从而影响文章结果的正确解释和应用。评阅中发现,有的文章样本量小于 20,作者却用 χ^2 检验,而没有用精确概率法;有的文章为计数资料,却采用 t 检验等错误。更重要的是,明确定义疗效指标是 RCT 统计分析、显著性检验和结果正确地解释及应用的基础。例如治疗组死亡率为 10%,对照组死亡率为 30%,两组死亡率差的倒数,即需治人数(number needed to treat, NNT)可用来作为疗效的指标^[8]。本例 $NNT = 1/(0.3 - 0.1) = 5$,含义为用此治疗方案,要避免 1 例死亡需治疗 5 个患者。另外,使用统一的疗效指标,如 NNT,可以直接比较不同治疗方案的疗效。但 414 篇文章中绝大部分未明确规定此类疗效指标。

2.3.8 副作用说明 任何一种治疗方案都有治疗作用和副作用两个方面。只有包括副作用的全面评价才能使读者获得安全性和有效性的综合信息。但本次评阅发现,281 篇(67.9%)文章没有对副作用进行任何说明。尽管部分文章提及有副作用,但并没有进一步对副作用类型、频率和严重程度进行描述。

2.3.9 评估结局依据 95% 以上的 RCT 使用客观指标作为测量结局的依据,这是减小偏性的重要措施之一,尤其是未用盲法时,应当尽量采用客观指标。

表 1 影响中医 RCT 文章质量的主要因素按年代分布 [篇(%)]

质量因素	81~85 年 (41 篇)	86~90 年 (110 篇)	91~96 年 (263 篇)	合计
描述随机方法	19(46.3)	19(17.3)	32(12.2)	70(16.9)
用非空白对照	39(95.1)	108(98.2)	252(95.8)	399(96.4)
有基线一致性比较	13(31.7)	44(40.0)	115(43.7)	172(41.5)
采用盲法	7(17.1)	22(20.0)	32(12.2)	61(14.7)
提及不依从	6(14.6)	16(14.5)	24(9.1)	46(11.1)
提及失访	10(24.4)	34(30.9)	31(11.8)	75(18.1)
提及更换措施	0(0)	1(0.9)	1(0.4)	2(0.5)
统计方法				
有统计量	8(19.5)	34(30.9)	107(40.7)	149(36.0)
有概率值	38(92.7)	92(83.6)	253(96.2)	383(92.5)
有副作用说明	14(34.1)	40(36.4)	79(30.0)	133(32.1)
评估结局				
依据客观指标	39(95.1)	107(97.3)	255(97.0)	401(96.9)

2.4 发表偏性问题 阳性结果的研究容易发表, 此即发表偏性⁽⁹⁾。这类偏性往往使读者对治疗效果形成错误的结论。本次评阅发现, 有统计学检验的 RCT 文章中, 阳性结果高达 86%; 对 45 篇降血脂治疗的 RCT 疗效进行漏斗图析⁽¹⁰⁾表明, 这类 RCT 发表中存在明显的发表偏性。这与国外新近的研究报道是一致的⁽¹¹⁾。

讨 论

自 1976~1996 年底中国医学学术期刊共发表评价中医药疗效的 RCT 约 10 000 篇, 是 MEDLINE 上可查到的评价传统医学 RCT 的 10 倍, 有关结果亟待总结和传播。同时, 现今中医药学术期刊中报道的 RCT 在设计、实施、分析和报告中存在一些问题。主要集中在随机分组方法描述, 盲法的使用, 不依从、失访和更换治疗措施的报告及处理, 疗效的定义和定量, 统计方法的正确使用, 样本量估计, 副作用的描述等。这些问题亦需要解决和改进, 否则, 花费大量的时间、人力和物力所收集的资料的价值就会受到质疑。

我们建议有计划地组织一批设计严格、样本量足够大、药物或疗法经严格选择、以主要疾病指标(如发病、死亡)来衡量疗效的 RCT。资源首先用于那些西医疗法不明显或花费昂贵、而中医措施极可能有效并且受益人数多的研究。有关中医疗效的 RCT 证据也应按照 Cochrane 协作的方法⁽²⁾不断地收集、总结和传播。

本次涉及质量评阅的 RCT 文章中有 90% 以西医诊断方法作为研究对象入选标准, 有 93% 以西医作为评估治疗结局的依据, 有 82% 的文章采用单方作为干预措施。这样的评估系统是目前较为推行的方法。但是否最合理和最科学, 是一个复杂的理论问题, 有待进一步研究和探索。

(本课题总负责为香港中文大学唐金陵教授。杨清、马胜琦、吴一舟、姬红、刘晓莉、李正懋同学参加资料收集工作, 本文承蒙北京医科大学临床流行病学专家曹家琪教授审阅, 本文部分

内容曾在 BMJ1999;319:160 刊出, 经该刊同意, 补充内容在本刊发表, 一并致谢)

参 考 文 献

1. Evidence-Based Medicine Working Group. Evidence-based medicine: A new approach to teaching the practice of medicine. JAMA 1991;268(17):2420—2425.
2. Chalmers I. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. Ann NY Acad Sci 1993;703:156—172.
3. Guyatt GH, Sackett DL, Cook DJ. Users' Guides to the Medical Literature II. How to use an article about therapy or prevention. JAMA 1993;270(21):2598—2601.
4. Guyatt GH, Sackett DL, Cook DJ. Users' Guides to the Medical Literature II. How to use an article about therapy or prevention. JAMA 1994;271(1):59—62.
5. 俞国培, 高思维, 李 怡, 等. 中药临床治疗试验质量的调查. 中国中西医结合杂志 1994;14(1):50—52.
6. Collins R, Peto R, Gray R, et al. Large-scale randomised evidence: trials and overviews. In: Weatherall DJ, Ledingham JGG, Warrel DA, eds. Oxford Textbook of Medicine 3rd ed. Oxford: Oxford University Press, 1996:21—32.
7. Newell DJ. Intention-to-treat analysis: Implications for quantitative and qualitative research. Int J Epidemiol 1992;21(5):837—841.
8. Cook RJ, Sackett DL. The number to treat: a clinically useful measure of treatment effect. BMJ 1995;310:452—454.
9. Eastbrook PJ, Berlin JA, Gopalan R, et al. Publication bias in clinical research. Lancet 1991;337:867—872.
10. Light RJ, Pillemer DB. Summing up. The science of reviewing research. Cambridge, MA: Harvard University Press, 1984:13.
11. Vickers A, Goyal N, Harland R, et al. Do certain countries produce only positive results? A systematic review of controlled trials. Controlled Clin Trials 1998;19:159—166.

(收稿:1998-10-09 修回:1998-12-30)

欢迎订阅本刊杂志

本刊 1988~1995 年各年合订本每册 50.00 元, 1996 年合订本每册 60.00 元, 1997、1998 年合订本每册 70.00 元(均含邮资); 1994~1996 年单行本(无 1995 年第 1、2 期)每本 3.90 元, 含邮资; 1997、1998 年单行本每本 4.90 元, 含邮资; 本刊自 1999 年第 1 期开始改为大 16 开版本, 订价每本 5.90 元, 含邮资。

本刊 1999 年基础理论研究特集即将出版发行, 每册订价 20.00 元(另加邮资 10%); 另外 1986~1998 年基础理论研究特集尚有余数。欲购者, 请直接向本社汇款订购。本社地址: 北京西苑操场 1 号, 中国中西医结合杂志社(邮编:100091)。开户银行: 北京工商银行海淀支行海淀分理处, 中国中西医结合杂志社, 帐号: 046098—24。务请写清收刊者详细地址、姓名及所在地邮政编码, 注明购某年合订本或特集及册数。