

DOI: 10.13288/j.11-2166/r.2015.18.009

国内应用 CONSORT 及 STRICTA 评价 RCT 的现状分析

温万鑫^{1,2}, 卢静敏², 陆丽明¹, 郭新峰^{*}

(1. 广东省中医院, 广东省广州市越秀区大德路 111 号, 510120; 2. 广州中医药大学第二临床医学院)

摘要 目的 对国内应用临床试验报告的统一标准 (CONSORT) 及其扩展版针刺临床试验干预措施报告标准 (STRICTA) 评价随机对照临床试验 (RCT) 的现状与问题进行系统总结。方法 计算机检索中国生物医学文献数据库 (CBM)、中国知网 (CNKI)、维普网 (VIP)、万方数据 (Wanfang) 数据库中应用 CONSORT 及 STRICTA 评价 RCT 的研究, 对其应用情况进行分析, 并对评价方法、报道方式、结果表示方法进行分析与总结。结果 共纳入 109 篇文献, 其中应用 2001 版 CONSORT 的 75 篇 (68.8%), 应用 2010 最新修订版的 31 篇 (28.4%); 评价中医药或中西医结合疗法 RCT 的研究有 74 篇 (67.9%); 由两位研究者独立进行评价, 如遇分歧通过协商或第三位研究者解决的占 81.7% (89 篇); 评价时采用二级评价 (报道, 未报道) 的有 72 篇 (66.1%), 采用三级评价 (充分报道, 部分报道, 未报道) 的有 6 篇 (5.5%); 使用列表方式或同时结合文字描述表示结果的有 75 篇 (68.8%)。结论 应用 CONSORT 声明及 STRICTA 评价 RCT 的研究越来越受关注, 临床研究报告质量呈上升趋势, 评价方法仍亟需提高与规范。

关键词 临床试验报告的统一标准; 针刺临床试验干预措施报告标准; 随机对照临床试验; 评价方法

设计科学和实施合理的临床随机对照试验 (randomized controlled trial, RCT) 可为系统评价、医疗的评估和决策提供高质量的“原始材料”, 正确地设计、执行和报告的 RCT 是评估干预性临床试验的金标准^[1]。但是, 随机化的对照研究也存在偏倚^[2], 许多 RCT 报告并不完善^[3], 没有严格方法学支撑的 RCT 将会产生偏倚的结果^[4-5], 缺乏完整报告的随机会对干预措施的评估带来偏倚^[6]。为了提高 RCT 的报告质量, 1996 年由临床试验学家、流行病学家、统计学家和生物医学杂志编辑等组成的国际小组制定了临床试验报告的统一标准 CONSORT (consolidated standards of reporting trials)^[7], 并分别于 2001 年^[8]和 2010 年^[9]进行了修订。CONSORT 的扩展版针刺临床试验干预措施报告标准 (standards for reporting interventions in clinical trials of acupuncture, STRICTA) 最早见于 2001 年^[10], 于 2008 年进行修订并正式成为 CONSORT 声明的官方扩展版^[11], 旨在改善针刺干预的临床试验报告规范。基于中医对疾病与治疗的认知有别于西医, 为了使得 CONSORT 更加适合中医药

的具体情况, 更好地提高中医药临床试验总体质量和报告质量, 国内中医药专家、方法学家等共同制订了中医药临床随机对照试验报告规范 (consolidated standards for reporting trials of traditional Chinese medicine, CONSORT for TCM) 以适用于中医药的药物临床试验^[12]。本研究系统地总结国内应用 CONSORT 及其扩展版 STRICTA 评价 RCT 的文献, 分析其现状和评价方法。

1 资料与方法

1.1 纳入标准

国内应用 CONSORT 及 STRICTA 评价 RCT 的文献。

1.2 排除标准

仅讨论或叙述 CONSORT 及 STRICTA 而非使用其评价 RCT 的文献; 无法获取全文的文献。

1.3 文献检索

计算机检索中国生物医学文献数据库 (CBM)、中国知网 (CNKI)、维普网 (VIP)、万方数据 (Wanfang)。检索时间为数据库收录起始时间至 2014 年 8 月 19 日。检索词主要包括“CONSORT”“STRICTA”, 并针对各数据库特点, 采取相应的检索策略。

* 通讯作者: guoxinfeng@139.com, 18922108539

1.4 文献筛选方法

检索出的文献条目导入 Endnote X7 (Thomson Scientific 公司), 由两位研究者独立对检出的条目进行筛选, 如遇分歧协商解决或征求第三方意见。首先进行查重, 然后阅读文题和摘要排除与研究不相关的研究, 最后下载全文, 逐篇阅读, 确定最终纳入的文献 (重复发表的研究, 保留数据最新最全的一篇)。文献筛选流程及结果见图 1。

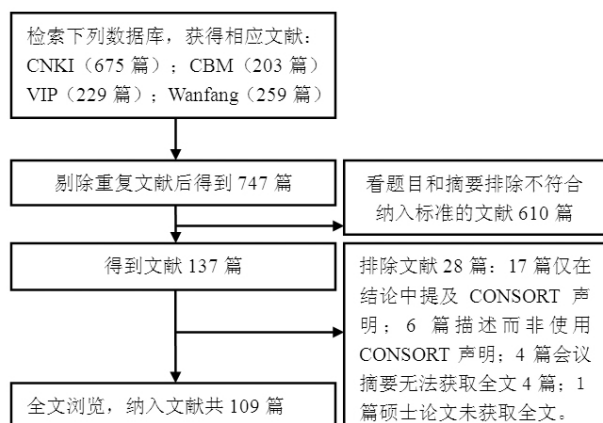


图 1 文献筛选流程及结果

1.5 数据处理与分析

最终纳入的文献由 Endnote X7 导出到 2007 版 Excel (微软公司), 两位研究者分别进行全文阅读与相关信息的抽取, 如遇分歧协商解决或征求第三方意见。并应用 Excel 进行合计与频数的描述性分析。

2 结果

2.1 国内 CONSORT 研究发表年份分布情况

图 2 示, 国内 CONSORT 研究的总趋势呈现曲

线上升, 2013 年达 23 个。纳入的 109 个研究中总共评价了 15975 个 RCT。

2.2 中西医应用 CONSORT 情况

采用 CONSORT 声明与 STRICTA 评价中医或中西医结合类 RCT 的研究有 74 篇 (67.9%), 评价西医 RCT 的有 18 篇 (16.5%), 另外 17 篇 (15.6%) 为评价某期刊或数据库, 或者某特定疾病。

2.3 CONSORT 不同版本应用情况

纳入的研究中, 使用 2001 年修订版^[8]及基于此版本的有 75 篇 (68.8%), 其中使用 CONSORT for TCM^[11]的有 12 篇, 在评价时加入了中医特色 (如有明确规范的中医诊断和中医证型的辨证标准、药物质量控制标准与方法、治则、中药模拟剂、中医证候疗效指标等) 的有 15 篇, 合并使用扩展版 STRICTA 的有 12 篇, 1 篇^[12]评价针刺 RCT 的研究未使用 STRICTA 清单。使用 2010 最新修订版^[8]及基于此版本的共有 31 篇 (28.4%): 其中 2 篇^[13-14]加入了中医特色条目 (如中医证型、组方依据、复方组成、制剂类型、质量控制、中医证候疗效指标、模拟剂等), 合并使用了 STRICTA 的有 14 篇。有 3 篇 (2.8%)^[15-17]只用了 STRICTA。所纳入的文献中, 使用有误的有 3 篇 (2.8%), 其中 2 篇^[16-17]描述使用的是 2010 版, 但文章结果所列条目则是 2001 版, 另 1 篇^[18]引文叙述使用 2001 版, 而文章所描述和结果列表显示的条目却是 2010 版。

2.4 评价方法

表 1 示, 所纳入研究中具体采用的评价方法。对评定者之间进行一致性检验的有 14 篇 (12.8%), 对原始材料不充分而联系作者加以确

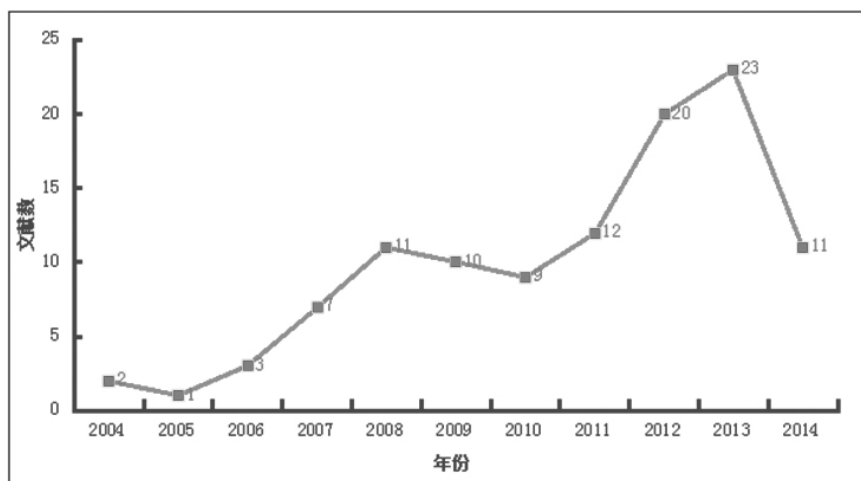


图 2 应用 CONSORT 及其扩展版评价 RCT 研究的年份分布情况

注: 2014 年的研究检索截止到 8 月份

定的有 8 篇 (7.3%)。在评价具体条目时采用二级评价 (报道, 未报道) 的有 72 篇 (66.1%), 采用三级评价 (充分报道, 部分报道, 未报道) 的有 6 篇 (5.5%); 评价时对条目进行评分的有 22 篇 (20.2%), 其中三级评分 (充分报道, 部分报道, 未报道) 的有 6 篇, 二级评分 (报道, 未报道) 的 16 篇。

表 1 应用 CONSORT 声明评价 RCT 的方法

评价方法	篇(%)
2 位评定者按清单的要求独立进行评价,如遇分歧,通过讨论达成共识	38(34.9%)
2 位评定者按清单的要求独立进行评价,如遇分歧,相互讨论解决或提交第 3 方协助解决	49(45.0%)
2 位评定者按清单的要求独立进行评价,如遇分歧,通过讨论达成共识,协调不一致时各自保留意见	1(0.9%)
2 位评定者按清单的要求独立进行评价,如遇分歧,通过讨论达成共识,协调不一致时由主要研究者决定	1(0.9%)
3 位评定者按清单的要求独立进行评价,如遇分歧,3 位研究员协商解决	1(0.9%)
1 位研究者单独评价,然后对数据进行复审,有差异返回原文核对	1(0.9%)
5 位评定者按清单的要求独立进行评价,如遇分歧,相互讨论,根据全体成员的 VAS 算术平均分确定	1(0.9%)
文章未提及评价方法	17(15.6%)

2.5 结果表示方法

使用列表方式或结合文字描述表示结果的有 75 篇 (68.8%), 应用条形图的有 3 篇 (2.8%)^[19-21], 对结果只进行文字描述的有 31 篇 (28.4%)。1 篇研究^[22]自行从 CONSORT 和 STRICTA 中抽取部分条目, 按照是否满足作者给出的分级标准来把文章的报告质量分为 3 级。还有 1 篇^[23]提及使用了 CONSORT 但却无具体的评价方法和评价结果。

2.6 其他评价方法

所纳入的研究中, 有 32 篇 (29.4%) 联合了 Jadad 评分量表^[24]评价方法学质量, 14 篇 (12.8%) 联合了 Cochrane 的偏倚风险评估工具^[25], 1 篇^[24] (0.9%) 联合了 GRADE 评价标准^[27]。

3 讨论

一个高质量的临床研究, 必须通过一份好的临床试验报告来体现^[28], CONSORT 声明就是为提高 RCT 报告质量而产生的^[6], 现在已经成为 RCT 国际报告标准^[29-31]。本研究纳入的 109 个研究中, 共评价了 15975 个 RCT, 结果显示, 国内研究者使用 CONSORT 评价 RCT 的研究逐年递增, 2013 年达到 23 个, 这趋势反映了 CONSORT 在国内的

关注度越来越高。国内已经有大量的中医药 RCT 报告发表, 试验报告质量参差不齐的现象已经引起了国内外的广泛关注^[31]。CONSORT 声明及其扩展版 STRICTA 作为临床科研方法学中的报告标准参考与推荐, 它的原理与方法具有普遍的指导意义, 不仅适用于现代医学, 同时也适用于中医药与中西医结合临床试验中, 其能促进中医药临床试验的发展, 提高其可靠性、客观性与论证强度, 使得中医药的研究成果为世人所公认, 促进中医药走向世界。由于中医药的特殊性, 其更加需要借助国际化的标准, 使其有效性被国际医学界所接受, 因此, 中医药界会更加关注 CONSORT 声明等系列标准化报告工具。本研究结果显示, 评价中医或中西医结合 RCT 的研究占 67.89% (74 篇), 而评价西医 RCT 的仅有 16.5% (18 篇), 表明我国中医药界非常注重提高临床试验方法和研究报告质量规范。国内学者也在其基础上制定出符合中医药特色的报告标准, 如《改进中药临床试验报告的参考条目》^[32], 《中医药临床随机对照试验报告规范》^[11]。毛兵等^[33]研究结果显示, 1999-2004 年中国 13 种中医药期刊 RCT 报告质量逐年提高但仍不理想, 随后通过评价《中国循证医学杂志》2001-2008 年发表的中医药 RCT 报告质量, 发现高于其前期研究^[34]。可见, 随着 CONSORT 不断完善, 中医药 RCT 报告质量在逐渐提高。

本文系统总结了国内所有应用 CONSORT 声明评价 RCT 的研究, 从纳入研究的评价方法可见, 由 2 位研究者独立进行评价, 如遇分歧通过协商或者第 3 位研究者解决的占 81.7% (89 篇), 可见大部分研究者也发现了单独 1 人评价的偏倚风险, 从而采取 2 位研究者独立进行评价, 如遇分歧通过协商或者第 3 位研究者解决的方法, 这样能提高评价的一致性, 减少测量偏倚。有些研究配合一致性检验, 或联系作者确定存在疑问的原始资料, 也有助于降低测量偏倚。

评价时采用二级评价 (报道, 未报道) 或采用三级评价 (充分报道, 部分报道, 未报道) 的共 78 篇 (71.6%), 此方法判断“未报道”是可行的, 而判断“充分报道或部分报道”则不妥, 因为“报告”的质量是由“报告”的完整性和准确性体现的^[7]。如果采用定量的方法测量, 则更能体现报告的完整性; 而对于报告的准确性, 如能把部分条目的细节描述具体, 如条目 3a、8a、8b、9、11a 等^[8], 把相应的试验设计、产生随机分配

序列的方法、随机方法的类型、隐藏和盲法都具体描述清楚,不仅有助于读者判断试验结果的准确性,而且有助于读者判断其是否符合系统评价的原材料,而加入备注一栏可助于解决此问题。如描述盲法时,有些文章提到此研究采用了双盲法,但细看文章内容,会发现有些研究的双盲对象为受试者和结局评价者,而真正意义上的双盲是研究者和受试者^[5],因此虽然其报道实施了盲法,但其是否被读者所认同,把具体内容描述出来,以供读者参考,可有助于读者判断结果的准确性。定量的方法和备注栏可能会使得评价稍显繁复,为了既保证评价结果的完整性和准确性,又不至于太繁复,推荐使用三级评价方法与备注栏相结合的报道方式。

纳入研究中,使用列表方式或同时结合文字描述表示结果的有 75 篇 (68.8%),用列表法将评价的内容用精辟的文字以表格的方式进行归纳和分类,可获得条理清晰、简单明了的效果。结合对其中重点条目进行文字的进一步阐述,能让读者对纳入研究有更详细和深入的了解。如果只采用文字描述,则使得阅读繁琐,效率低下,且不能很好地抓住重点。

本研究发现,有 3 篇文献存在使用错误,2 篇评价针刺 RCT 没应用相关的 STRICTA 清单,表明国内研究者对 CONSORT 的认识还有待提升与普及。本研究只纳入了国内 4 大数据库所收录的文献,不免遗漏其他数据库的相关研究。

综上,国内应用 CONSORT 等系列随机对照试验标准化报告工具评价 RCT 的研究越来越多,临床研究报告质量呈上升趋势,但评价的过程和方法学细节尚待进一步提高。我们推荐,由两位研究者独立进行评价,如遇分歧通过协商或者第三位研究者解决的评价方法;使用三级评价与备注栏相结合的报道方式;使用列表方式同时结合文字描述表示结果的方法等。以助研究者能便捷标准地使用,读者能更准确地评价临床试验报告的质量。

参考文献

- [1] Byar DP, Simon RM, Friedewald WT, et al. Randomized clinical trials: perspectives on some recent ideas [J]. N Engl J Med, 1976, 295(2): 74-80.
- [2] Schulz KF, Chalmers I, Grimes DA, et al. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals [J]. JAMA, 1994, 272(2): 125-128.
- [3] Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials [J]. JAMA, 2001, 285(15): 1987-1991.
- [4] Schulz KF, Chalmers I, Hayes RJ, et al. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials [J]. JAMA, 1995, 273(5): 408-412.
- [5] Jüni P, Altman DG, Egger M. Systematic reviews in health care: assessing the quality of controlled clinical trials [J]. BMJ, 2001, 323(7303): 42-46.
- [6] Moher D, Pham B, Jones A, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? [J]. Lancet, 1998, 352(9128): 609-613.
- [7] Begg C, Cho M, Eastwood S, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement [J]. JAMA, 1996, 276(8): 637-639.
- [8] Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials [J]. BMJ, 2010, 340: c869. doi:10.1136/bmj.c869
- [9] MacPherson H, White A, Cummings M, et al. Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendation [J]. Complement Ther Med, 2001, 9(4): 246-249.
- [10] MacPherson H, Altman DG, Hammerschlag R, et al. Revised STAndards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA): extending the CONSORT statement [J]. Acupunct Med, 2010, 28(2): 83-93.
- [11] 吴泰相,李幼平,卞兆祥. 中医药临床随机对照试验报告规范(征求意见稿) [J]. 中国循证医学杂志, 2007, 7(8): 601-605.
- [12] 赵相锋. 国内针刺治疗近视眼疗效的系统评价 [D]. 天津: 天津医科大学, 2011.
- [13] 贾永亮,黄芳艺,张时开,等. 中成药治疗冠心病随机对照试验方法学质量评价 [J]. 中国中西医结合杂志, 2014, 32(4): 560-568.
- [14] 高振,刘莹莹,李凤森,等. 穴位贴敷治疗慢性阻塞性肺疾病稳定期随机对照试验质量评价 [J]. 中国中医药信息杂志, 2014, 21(4): 17-21.
- [15] 孙亦乐. 国内针刺治疗中风痉挛性偏瘫随机对照试验的文献质量探究 [D]. 北京: 北京中医药大学, 2013.
- [16] 王晓庆,段培蓓,张晓琴,等. 我国关于肿瘤化疗患者预期性恶心呕吐干预性研究文献方法学质量评价 [J]. 护理学报, 2014, 21(14): 1-4.
- [17] 王凤蕊. 基于循证医学方法系统评价针灸治疗实证眩晕的临床效应 [D]. 杭州: 浙江中医药大学, 2013.
- [18] 陈静雅,潘华山,冯毅舫. 基于 CONSORT 和 STRICTA 评价针灸治疗肠易激综合征临床随机对照试验报告质量 [J]. 辽宁中医药大学学报, 2013, 15(3): 143-146.

- [9] 陈军, 索娟, 曾晓敏, 等. 中医药防治胆石症随机对照试验的质量评价 [J]. 中国循证医学杂志, 2008, 8(5): 370-374.
- [20] 赵刚, 任冷澜, 刘雅莉, 等. 中文骨科期刊相关随机对照试验文献的质量评价 [J]. 甘肃科技, 2010, 26(5): 45-47.
- [21] 屈建峰. 针灸治疗肩周炎随机对照文献质量评价研究 [D]. 北京: 中国中医科学院, 2011.
- [22] 刘雪梅, 张鸣明, 杜亮. 针刺减肥的随机对照试验方法学和报告质量评估 [J]. 中国医学科学院学报, 2004, 26(2): 192-194.
- [23] 王晓东, 贾洪, 闫军. 药物治疗麻风性结节性红斑随机对照临床试验的评价 [J]. 中国麻风皮肤病杂志, 2010, 26(5): 327-330.
- [24] Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? [J]. Control Clin Trials, 1996, 17(1): 1-12.
- [25] Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [EB/OL]. (2011-03). www.cochrane-handbook.org.
- [26] 何梅, 杨柳, 马宗斌, 等. 复方甘草酸苷治疗荨麻疹的随机对照试验报告质量评价 [J]. 中国医院药学杂志, 2012, 32(19): 1564-1571.
- [27] 刘为民, 刘保延, 何丽云, 等. 引入 GRADE 体系制定中医药国际标准的思考与实践探析 [J]. 亚太传统医药, 2011, 7(1): 1-5.
- [28] 温泽淮, 赖世隆, 梁伟雄, 等. 进一步改善中医药临床试验报告的质量 [J]. 中国中西医结合杂志, 2004, 24(11): 1022-1025.
- [29] Moher D, Schulz KF, Altman DG, et al. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials [J]. Ann Intern Med, 2001, 134(8): 657-662.
- [30] Moher D, Jones A, Lepage L, et al. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation [J]. JAMA, 2001, 285(15): 1992-1995.
- [31] 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.
- [32] 费宇彤, 刘建平. 改进中药临床试验报告的参考条目 [J]. 中西医结合学报, 2008, 6(3): 233-238.
- [33] 毛兵, 王刚, 樊涛, 等. 中医药随机对照试验报告质量评价 [J]. 中国循证医学杂志, 2007, 7(12): 880-887.
- [34] 闵婕, 毛兵, 刘艾然. 《中国循证医学杂志》发表的中医药随机对照试验质量评价 [J]. 中国循证医学杂志, 2010, 10(3): 362-366.
- [35] 王家良. 临床流行病学-临床科研设计、测量与评价 [M]. 3 版. 上海: 上海科学技术出版社, 2009: 72.

Analysis of the Current Status of RCTs Evaluation with CONSORT Statement and STRICTA in China

WEN Wanxin^{1,2}, LU Jingmin², LU Liming¹, GUO Xinfeng¹

(1. Guangdong Hospital of Traditional Chinese Medicine, Guangzhou, Guangdong Province, 510120; 2. Second Clinical School of Guangzhou University of Chinese Medicine)

ABSTRACT Objective To make systemic summary on the status and problems of randomized controlled trials (RCTs) evaluation with Consolidated Standards of Reporting Trials (CONSORT) statement, which is the unified standard of clinical trial reports, and Standards for Reporting Interventions in Clinical Trials of Acupuncture (STRICTA), which is expanded standards for reporting interventions in controlled trials of acupuncture interiorly, in China.

Methods Studies on RCTs evaluated with CONSORT and STRICTA were searched in China Biomedical Bibliographic Database (CBM), China National Knowledge Infrastructure (CNKI), VIP Database, and Wanfang Database. The application of CONSORT and STRICTA was analyzed, and the evaluation methods, reporting styles, and ways of showing results were summarized. **Results** 109 papers were included in. Among them, CONSORT of 2001 edition was applied in 75 papers (68.8%), CONSORT of 2010 (latest revised) edition was applied in 31 papers (28.4%); 74 papers (67.9%) were about evaluation on Chinese Medicine RCTs or integrated Chinese Medicine and Western Medicine treatments RCTs; 81.7% of evaluations (89 papers) were made by two researchers independently, and different opinions were solved by consultation or a third researcher; two-level evaluation (reported, unreported) was adopted in 72 papers (66.1%), and three-level evaluation (fully reported, partly reported, unreported) was used in 6 papers (5.5%); and 75 papers (68.8%) showed results by list or list combined with text description.

Conclusion Researches on RCT evaluation by CONSORT and STRICTA have attracted more and more attention. The quality of clinical research reports is getting better, but the evaluation methods need to be improved and regulated urgently.

Keywords Consolidated Standards of Reporting Trials statement; Standards for Reporting Interventions in Clinical Trials of Acupuncture; report quality; randomized controlled trials; evaluation method

(收稿日期: 2015-01-21; 修回日期: 2015-05-10)

编辑: 焦爽