

• 方法学与临床评价 •

中医药治疗失眠临床试验中结局指标的选择与报告

曹卉娟 刘建平

(北京中医药大学循证医学中心, 北京市朝阳区北三环东路 11 号, 100029)

[摘要] 针对当前中医药临床试验关于结局测量指标应用不统一、与国际不接轨等迫切需要改善的现状, 以中医药治疗失眠为例剖析了疗效评价中涉及结局指标选择的相关注意事项, 给出了具体的国际化标准及规范要求。结合具体的实例分析, 以求改进国内目前对疗效指标选择应用较为混乱的局面, 使中医药临床试验在结局测量和结果报告方面逐步达到国际统一规范, 促进中医药临床研究质量的提高。

[关键词] 中医药疗法; 失眠; 结局测量指标; 临床试验报告

Selection and Reporting of Outcome Measures in Clinical Trials on Insomnia Treated with TCM Therapy

CAO Huijuan, LIU Jianping

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

ABSTRACT Because of the disunity use of the outcome measures in traditional Chinese medicinal clinical trials, we explain the application of outcome measures in clinical trials on insomnia, provide the international criteria and guideline on outcome reporting. According to analysis of the concrete cases, we in order to improve the misuse of outcome measures of insomnia, to make the selection and reporting of outcome measures follow with the international practices and, to accelerate the development of traditional Chinese medicine.

Key Words TCM therapy; Insomnia; Outcome measures; Report of traditional Chinese medicinal clinical trial

一项干预措施对于目标疾病的治疗是否有效, 往往反映在疾病的结局测量指标的正确选择以及指标的改善方面。因此, 选择恰当的结局测量指标就成为评价干预措施有效性中十分重要的因素。目前, 中医药临床试验关于结局测量指标及疗效评价方面存在诸多问题, 体现在“标准”过多而且判断尺度不同, 无论是标准的使用还是结局指标的选择都存在与国际标准不接轨、应用不恰当的弊端。本文以中医药治疗失眠为例剖析结局评价中存在的问题。

1 临床试验中的结局测量指标的选择

临床试验在设计阶段必须要考虑的一个重要方面就是如何选择能够体现疗效的结局指标。其依据应当是按照干预措施作用的特点和可能的作用机理来加以选择。例如, Ⅱ期临床试验着重于效力和机制的探讨, 经常会采用一些替代指标(尤其是实验室检测的指标), 而Ⅲ期临床试验是要解决干预措施的效果和安全性问题, 因此, 通常会选择一些临床结局指标。而中医

药临床试验的判效标准中更多是属于临床结局指标, 即能够体现患者相关的结局指标。

结局测量指标按照指标资料类型分类, 可以分为定性指标、定量指标和时间序列指标^[1]。定性指标一般为二分类变量或等级变量, 如有效与无效, 阳性与阴性。定量指标一般为精确测量得到的连续变量数值, 如血压值、疼痛视觉模拟评分(VAS)分值等。时间序列指标如生存率、复发率等。临床试验中另外一种结局测量指标还可分为替代指标和终点指标^[2]。近年来, 循证医学强调疾病的终点结局, 也就是疾病的最终结局, 是患者及研究者对于疾病的发展最为关注的临床事件, 如死亡、骨折、痊愈等。替代指标是一种能代替重要临床结局的实验室或影像学测量指标, 如血脂水平、肿瘤大小等等。指标的准确选择应当根据干预措施的特点、治疗目标和疾病性质综合考虑。

2 国内常用的失眠疗效评价标准和结局测量指标以及存在的相关问题

失眠是一种常见的疾病, 中医药治疗失眠常见的特色疗法有中药、针刺、推拿、按摩等。有关中医药或中西医结合治疗失眠的临床研究报道不胜枚举, 以针刺治疗失眠为主题对国内的期刊数据库进行检索, 得到

基金项目: 国家重点基础研究发展计划(“973”计划)资助项目(2006CB504602)、科技部国际合作项目、中澳中医临床疗效评价研究项目(2009DFA31460)

通讯作者: 刘建平, jianping-1@hotmail.com, (010) 64286757

了数以百计的临床试验报告,包括 1 篇中文发表的系统评价^[3]。综合这些临床试验的结局指标,按照出现频率排序,大体可以分为 5 类。第一类是对患者睡眠情况进行问卷和量表的评分,包括匹兹堡睡眠质量评分(PSQI)、失眠严重程度评分(ISI)、睡眠功能障碍等级量表(SDRS)、晨起调查问卷(MQ)、自评焦虑等级量表(SAS)等评分原则,并以总分或各项分值作为结局指标。第二类是依据《中药新药临床研究指导原则》^[4],将疗效分为 4 个等级,即临床治愈:睡眠时间恢复正常或夜间睡眠时间在 6h 以上,睡眠深沉,醒后精神充沛;显效:症状明显好转,睡眠时间增加 3h 以上,睡眠深度增加;有效:症状减轻,睡眠时间较前增加不足 3h;无效:治疗后失眠无明显改善或加重者。第三类是根据《中医病证诊断疗效标准》^[5]将患者分为 3 个等级,即治愈:睡眠正常,伴有症状消失,入睡时间 < 30min, 24h 总睡眠时间 > 8h;有效:睡眠时间延长,伴有症状改善,入睡时间 < 30min, 24h 总睡眠时间 > 4h;无效:症状无改变,入睡时间 > 30min, 24h 睡眠 < 4h。第四类是采用国际统一睡眠效率值(简称睡眠率),睡眠率 = 实际入睡时间/上床至起床总时间 × 100%,根据睡眠率将患者分为 4 个等级,痊愈:症状消失,睡眠率 75% 以上,停用安眠药;显效:症状缓解,睡眠率 65% 以上,停用安眠药;有效:症状改善,睡眠率 55% 以上,基本停用安眠药或减去 3/4 量;无效:症状如前,睡眠率 40% 以下,靠安眠药维持。第五类是疗效标准仅为等级标准,但其等级标准并非明确统一的标准,或其标准仅与患者自觉症状改善有关,无明确的参考数值。

就这 5 类疗效评价标准来看,第二~第五类结局指标虽然具体分类方法不同,但都属于等级变量即定性指标,从内容上看,每一等级的定义都由几项疗效指标构成,因此属于综合性判效指标,即涵盖症状、体征、睡眠时间、用药情况等。以《中药新药临床研究指导原则》的 4 个等级为例,对于“临床治愈”的定义包括“睡眠时间恢复正常或夜间睡眠时间在 6h 以上”、“睡眠深沉”、“醒后精神充沛”3 项,这 3 项之间的逻辑关系并没有交代清楚,“睡眠时间恢复正常”的具体水平含混,缺乏定义;使研究者在实际判断中存在不确定性,导致不同研究者之间在理解和判断上存在不一致的情况,因此,整体疗效评价的客观性欠佳。尤其第五类结局指标,不但多为综合性指标,且无统一的等级分化标准,甚至为研究者自拟等级,使得研究结果无法得到客观准确的评价,研究的内部真实性值得怀疑。在研究报告方面,除了统计分析方法交代不清外,有些研究为了获

得较高阳性结果,将成文的标准进行改动后报告结果,如将“睡眠率 55% 以上为有效”改为“睡眠率 50% 以上为有效”^[6],将“睡眠时间增加 3h 以上为显效”改为“睡眠时间增加 2h 以上为显效”^[7],使得原本就含糊不清的分类标准更加杂乱不统一。

第一类评分作为连续变量的定量指标,属于国际常用指标,以匹兹堡睡眠质量评分^[8]为例,该评分总分涵盖睡眠质量、入睡时间、睡眠时间、睡眠效率、睡眠紊乱、安眠药物使用、日间活动功能 7 项内容,能较好地反映干预措施对于失眠程度的改善,然而很多研究在报告这一结局指标时或仅报告总分^[9],或未报告总分^[10],或报告治疗后的评分而未报告治疗前的评分^[11],而且绝大多数研究都未报告结局指标的统计分析方法。

3 推荐使用的治疗失眠的结局测量指标

在 Cochrane 协作网上检索针刺治疗失眠的研究获得 1 篇英文发表的系统评价^[12],文中结局指标分为两种:主要结局,即患者自述对于失眠改善情况,即对睡眠质量的评价,以分值出现,相当于患者报告的临床结局(patient report outcome, PRO)。次要结局,包括睡眠因素,如入睡时间、总睡眠时间;与睡眠相关的标准量表,如匹兹堡睡眠质量评分;日间活动情况,可以通过一些评分来测定;以及生命质量、不良事件发生的频率。

此外,我们还检索了截止到 2008 年 12 月的国外期刊发表的临床试验报告,关于结局测量指标多采用连续变量为主的定量指标,如治疗前后入睡时间、睡眠时间的均值和标准差^[13]、治疗前后匹兹堡睡眠质量评分和生活质量评分的均值和标准差^[14],这一类指标能较客观而清晰地报告疾病相关症状在治疗前后的变化情况,而且由于标准统一、定义明确,能够对结果进行解释和分析,不同研究之间的结果也可以比较,因此在系统评价/Meta 分析时能更好地进行资料综合利用。从当前国外发表的文献可见,国际上在报告某种干预措施治疗失眠的定量研究疗效评价和结局测量指标的时候,多采用定量指标,而较少使用定性指标,几乎没有等级分类的指标。而作为时间序列指标,睡眠率也可以直接作为结局测量指标来报告,其结果无需进行额外的计算来进行等级划分,可以直接报告治疗前后睡眠率的变化^[11]。

因此,在做失眠相关的临床研究时,选用的结局测量指标应以上述国际常用指标为主,如表 1 所示。

4 临床试验中涉及疗效指标报告的注意事项

对于干预措施疗效评价的临床试验,尤其是随机

对照试验,其报告应该遵循国际 CONSORT (Consolidated Standards of Reporting Trials) 标准,高质量的试

表 1 临床评价研究中推荐使用的失眠相关结局测量指标	
测量结局	对应的结局测量指标
主要结局	患者对失眠改善的评价
次要结局	患者对睡眠质量等改善情况的评价 (PRO)
	入睡时间、总睡眠时间、总清醒时间、入睡后醒来的情况、夜间醒来或早醒的情况、睡眠率*
	睡眠相关的标准化量表
	匹兹堡睡眠质量评分
	斯坦福瞌睡量表 (Stanford Sleepiness Scale)、波沃斯嗜睡量表 (Epworth Sleepiness Scale)
	日间活动情况
不良事件	生活质量
	SF-36 量表
不良事件	不良事件发生率*

注: * 标注为等级指标,其余均为连续变量指标

验报告可以真实反映临床试验结果的真实性和应用性^[15],在“CONSORT 声明”中关于临床试验结局报告的部分有 7 条规范要求^[16],在试验设计、实施良好的前提下,对于结局的报告尤为关键,以失眠治疗为例,试验报告应该在结果部分包含以下内容。

- ①受试对象流程图。说明受试对象在试验各个阶段的数目和脱落失访情况。
- ②对象纳入的持续时间。说明试验从纳入第一例病例到最后一例的时间段及随访的情况。
- ③基线资料。国内大部分文献都把基线资料放在研究方法的部分交代,实际上基线资料的每一项,如年龄、性别、病程、治疗前各项指标测量值等都应该在结果部分详细的列表说明,不能简单地以“两组基线资料无统计学差异”的交代略过。
- ④分析的病例数。这里同时要说明结果分析是否采用了意向性分析(intention-to-treat, ITT)的方法,即对脱落失访病例结局测量值的处理方法。根据 Cochrane 方法中关于脱落失访病例数据的处理方法的指南^[17],采用 ITT 的方法能更好地体现所有分配入组的各个纳入患者的指标,较之全数据集(PPA)的统计分析结果更能说明问题。ITT 的处理方法有 5 种,处理连续变量时,将缺失患者最后一次测量到的数据作为结局指标纳入统计。如果是定性指标,如等级分类变量数据,则还可以有最优结果分析、最差结果分析、极端最优、极端最差结果分析 4 种 ITT 分析方法,更能说明试验组与对照组差异的大小。以表 2 数据为例,我们统计治疗组与对照组关于不良事件发生率的差异比较。运用 PPA 数据集的统计分析,即不考虑脱落例数

对两组不良事件发生率进行比较,结果两组在不良事件发生率上差异有统计学意义($P < 0.0001$);而运用 ITT 中极端最差结果分析方法时,两组治疗总例数均为 300 例,其中治疗组不良事件发生例数为 50 例(30+20),对照组不良事件发生例数为 70 例,在不良事件发生率比较上,两组差异无统计学意义($P = 0.096$)。由此看出,这一例试验采用两种分析方法的结果大相径庭,在数据分析阶段,ITT 分析是一种很重要的处理缺失数据的方法。

表 2 ITT 方法运用实例数据(例)				
组别	有不良事件发生	无不良事件发生	脱落例数	总例数
治疗组	30	250	20	300
对照组	70	200	30	300

- ⑤结局和效应大小。报告每一项主、次要结局,需要交代结果指标测量的时间点、测量方法及质量控制,继而给出原始数据及分析结果,如匹兹堡睡眠质量评分 7 项各自的均值、标准差及总分的均值、标准差,并给出治疗前后比较的统计分析结果(P 值)。需要强调的是,在研究方法部分应交代清楚对每一项测量值所使用的统计分析方法及统计软件。
 - ⑥亚组或校正分析。如果事先说明了亚组和校正因素,则应对相应的内容进行附加的资料分析。
 - ⑦不良事件。报告试验各组的不良事件、副作用或药物不良反应。
- 鉴于中医药自身的特点,国内外专家在 CONSORT 条目的基础上,制定了相关的规范声明,以期提高中医药随机对照试验的报告质量,如针灸临床试验报告标准^[18]、国际草药 CONSORT 声明^[19],中药临床试验干预措施报告建议^[20]等。可以看到,在国际规范中,更多强调的是 PRO,因为 PRO 是直接来自于患者对于疾病改善情况和自身状态的评价,包括患者描述的功能状况或客观症状性指标和与健康相关的生存质量,属于中间指标,其注重整体的思想更接近于中医学思想^[21]。
- 综上所述,有三点问题是当前需要迫切关注和加以改善的,一是要运用符合国际化标准的疗效指标来进行结局测量与评价,针对中医干预的特点、研究目标及疾病本身特点,参照国际规范决定选择合适的指标。二是要客观真实地报告结果,不要选择性地报告疗效测量和分析结果,以减少偏倚。三是在报告临床试验时,要采用 CONSORT 等规范化的报告标准。总之,在进行临床试验报告时,应尽量遵循国际规范,重视中间指标和 PRO 的运用,提高临床试验报告的质量。

参考文献

[1] 刘建平. 临床试验效应的表达——指标及其意义[J]. 中西医结合杂志, 2004, 24(8) : 677—679.

[2] 徐小国, 蒋萌. 替代指标在临床试验中的应用与局限性[J]. 中国临床药理学与治疗学, 2006, 11(10) : 1194—1197.

[3] 李宁, 吴滨, 王成伟, 等. 针灸治疗失眠症随机对照试验的系统评价[J]. 中国针灸, 2005, 25(1) : 7—10.

[4] 中华人民共和国卫生部. 中药新药临床研究指导原则[M]. 1993: 186—188.

[5] 国家中医药管理局. 中医病证诊断疗效标准[S]. 南京: 南京大学出版社, 1994: 175.

[6] 陆卫卫. 针刺内关配合七叶神安片治疗顽固性失眠 35 例[J]. 南京中医药大学学报, 1998, 14(3) : 164.

[7] 刘琴. 针灸配合耳穴治疗不寐 36 例[J]. 吉林中医药, 2000, 20(4) : 63.

[8] Tsai PS, Wang SY, Wang MY, et al. Psychometric evaluation of the Chinese version of the Pittsburgh Sleep Quality Index (CPSQI) in primary insomnia and control subjects[J]. Qual Life Res, 2005, 14(8) : 1943—1952.

[9] 刘健红, 黄坚红, 陈秀慧. 针刺治疗中风后失眠 32 例疗效观察[J]. 国际医药卫生导报, 2006, 12(16) : 107—108.

[10] 宣雅波, 郭静, 王麟鹏, 等. 针刺对原发性失眠患者睡眠质量的影响: 随机对照研究[J]. 中国针灸, 2007, 27(12) : 886—888.

[11] 翁明, 廖海清. 电针治疗老年性失眠疗效量表分析[J]. 针灸临床杂志, 2007, 23(5) : 33—34.

[12] Cheuk DK, Yeung WF, Chung KF, et al. Acupuncture for insomnia[J]. Cochrane Database Syst Rev, 2007, 18(3) : CD005472.

[13] Suen LK, Wong TK, Leung AW. Effectiveness of auricular therapy on sleep promotion in the elderly[J].

Am J Chin Med, 2002, 30(4) : 429—449.

[14] Tsay SL, Rong JR, Lin PF. Acupoints massage in improving the quality of sleep and quality of life in patients with end-stage renal disease[J]. J Adv Nurs, 2003, 42(2) : 134—142.

[15] 刘建平. 循证中医药临床研究方法学[M]. 北京: 人民卫生出版社, 2006: 67—68.

[16] 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]. Lancet, 2001, 357(9263) : 1191—1194.

[17] The Cochrane Collaboration. Further issues in meta-analysis: intention to treat issues [EB/OL]. [2009-04-20]. <http://www.cochrane-net.org/openlearning/html/mod14-4.htm>.

[18] MacPherson H, White A, Cummings M, et al. Standards for reporting interventions in controlled trials of acupuncture: the STRICTA recommendations [J]. Acupunct Med, 2002, 20(1) : 22—25.

[19] Gagnier JJ, Boon H, Rochon P, et al. Reporting randomized, controlled trials of herbal interventions: an elaborated CONSORT statement [J]. Ann Intern Med, 2006, 144(5) : 364—367.

[20] 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.

[21] 金香兰, 张允岭, 韩振蕴, 等. 关于引入 PRO 理念评价中医药治疗轻度痴呆疗效的思考[J]. 天津中医药, 2008, 25(1) : 29—31.

(收稿日期: 2009—04—29; 修回日期: 2010—02—23)

[编辑: 崔京艳]

(上接第 404 页)

夜交藤等。其中尤其喜合用合欢花、合欢皮, 二者一气一血, 花轻清走上, 皮和血下行, 既入肝经又入心经, 既可解郁又可安神, 治疗虚烦不安, 是调心安神之佳品。《神农本草经》云合欢皮“主安五脏, 和心志, 令人欢乐无忧”; 久泻伤阴, 夜交藤归心、肝经, 味甘能补, 有滋阴养血、养心安神之效, 常与合欢皮相须为用, 《饮片新参》云其“养肝肾、止虚汗, 安神催眠”。此外, 还应针对引起心神失常的病因进行调理治疗, 如由于心阴亏虚, 而出现心烦失眠、多汗等症状时, 可选用百合、生地黄、浮小麦等养心安神。心脾两虚而见不易入睡、多梦易醒、心悸健忘、神疲食少等, 可选用白术、甘草、当归、黄芪、五味子益气养心。腹痛明显者可予延胡索、白芍理气柔肝止痛; 腹胀明显者予枳壳、木香、荔枝核理气消胀; 腹泻明显者用健脾利湿, 收涩止泻之品, 如芡实、莲子、益智仁。治疗本病应避免应用磁石、朱砂等金石重镇安神之品, 以防止清阳难升, 泄泻更甚。

3 典型病例

李某, 男, 39 岁。腹痛腹泻反复发作 5 年, 腹痛隐隐, 部位不固定, 腹痛即泻, 每日 3~10 次, 多为水样便, 每于饮食不慎及精神紧张时加重, 肠鸣腹胀, 食少、面色萎黄, 肢体倦怠, 失眠多梦, 心烦, 纳差, 舌淡红、苔白微腻, 脉弦细。电子肠镜未见明显异常, 大便常规示水样便、色黄, 镜检见有未消化食物。西医诊断为腹泻型 IBS, 中医诊断为泄泻, 辨证为肝脾失调、心神失养, 治以疏肝运脾, 调心安神。处方: 隔山消 15g, 柴胡 12g, 白芍 20g, 当归 12g, 白术 10g, 茯苓 15g, 枳实 12g, 八月札 15g, 佩兰 12g, 合欢皮 20g, 合欢花 12g, 夜交藤 20g, (炒) 谷芽、(炒) 麦芽各 20g, 砂仁 6g, 石榴皮 15g, 芡实 12g, 葛根 20g, 防风 6g, 服药 7 剂, 复诊时大便每日一二次, 基本成形, 腹痛缓解, 睡眠好转。后用上方加减调理 2 周而病愈, 随访 1 年, 未复发。

(收稿日期: 2009—10—20)

[编辑: 洪涛]