

中医药临床试验的方法学问题与挑战：循证医学的观点

刘建平

(北京中医药大学基础医学院循证中医药临床研究与评价中心, 北京 100029)

[摘要] 运用循证医学观点分析了中医药临床试验及其评价的现状, 论述了随机、对照、盲法在中医药临床试验及其疗效评价研究中的应用条件, 并阐明了在中医药临床疗效评价研究中如何估算样本量、选择结局评价指标及临床试验依从性与意向性治疗分析。将循证医学应用于中医药临床试验, 有助于提高中医药临床诊疗与研究水平、促进国际间的合作与交流。建立循证的中医药临床疗效评价体系, 有利于客观、科学地评价中医药的有效性与安全性。

[关键词] 医学, 中国传统; 循证医学; 临床试验; 方法学

[中图分类号] R2-05 [文献标识码] A [文章编号] 1672-1977(2006)01-0001-06

Methodological quality assessment of clinical trials in traditional Chinese medicine: the principles of evidence-based medicine

Jian-Ping LIU

(Evidence-based Chinese Medicine Center for Clinical Research and Evaluation, School of Preclinical Medicine, Beijing University of Chinese Medicine, Beijing 100029, China)

ABSTRACT The quality of clinical trials in traditional Chinese medicine (TCM) was analyzed with respect to their methodology and outcome assessment according to the concepts of evidence-based medicine (EBM). It was suggested that the essential methodological principles of EBM should be applied in clinical trials of TCM, including randomization, control and blindness. The sample size estimation, optimal parameter selection for outcome assessment, compliance and intention-to-treat analysis are all important aspects of randomized controlled trials. To bring the benefits of EBM into clinical trials will improve the quality of research in TCM and also promote international co-operation and communications. It is necessary to establish an assessment system of clinical outcome for TCM on the basis of EBM in order to evaluate the efficacy and safety of TCM objectively and scientifically.

KEY WORDS medicine, traditional Chinese; evidence-based medicine; clinical trials; methodology

Zhong Xi Yi Jie He Xue Bao/ J Chin Integr Med, 2006, 4(1):1-6 www.jcimjournal.com

1 中医药的临床评价

中医药之所以能够长期存在, 在于其防治疾病的有效性。尽管许多中医药疗法在中国和其他国家被广泛应用, 但由于其理论体系和治疗原理无法用西医的病理生理机制加以阐释, 因此尚未得到西方医学界的普遍认可。其主要问题在于: (1) 中药制剂, 尤其是复方制剂, 其有效成分的基础研究还不够深入, 不能提供诸如主要成分含量、理化特性、生物

学活性和作用靶点等科学实验证据; (2) 中医辨证论治强调治疗的个体化, 因此难以做到标准化, 限制了中医药在更大范围内的应用。

鉴于西方国家越来越多的民众在使用传统医药尤其是中医药进行治疗, 许多国家已开始重视这些补充和替代医学疗法。如美国国立卫生研究院的补充和替代医学研究中心就获得了政府大量的研究经费, 用于传统医药的临床评价研究。中医药要走向世界, 就必须用国际公认的标准去衡量其疗效, 以求

提供客观、科学的证据。

中医药的临床评价应当是在中医理论指导下,对中医临床诊断和治疗加以验证,证明其有效性与安全性。中医临床评价是临床医师、患者以及卫生行政部门共同关心的问题,它关系到进一步的临床决策和卫生资源的合理分配,也关系到医疗质量的改进和医生业务素质提高等诸多方面。传统的临床评价多依赖中医专家和医师在临证实践过程中对个案病例或系列病例的经验总结,缺乏严格设计的前瞻性试验研究。其明显的不足在于偏倚难以控制,结果经不起重复。因此,临床经验难以得到提升,好的疗法得不到进一步的推广应用。循证医学提倡的证据是指通过前瞻性随机对照试验所获得的结果,是针对某一疾病或病症采用某一干预措施对随机选择的病例进行试验干预后所观察到的客观效应。循证医学的另一种级别更高的证据则是对单个随机研究证据进行系统、全面的鉴定与评价,即系统性评价或荟萃分析。因此,中医的证候研究和疗效评价均可采用循证医学的研究方法。

传统中医的辨证论治虽然也是一种原始的寻证方法,但中医的“证”与循证医学的“证”具有不同的范畴和内涵。中医的“证”是指对疾病或亚健康状态通过望、闻、问、切等观察手段获得的表象,“证候”则是对这种表象及其动态变化的综合表述。因此,这种证据是不全面的,常因观察者不同而相异,易于发生偏倚。循证医学的“证”是指通过严格设计的人体试验研究所获得的客观、真实的结果,是经得起验证和重复的。因此,将循证医学方法应用于中医的临床疗效评价,能够得出客观、科学、系统的研究结果。然而,作为一门新兴的方法学,循证医学也不是万能的,对于中医辨证论治提倡的个体化治疗(individualized treatment),随机对照临床试验就有其局限性。如何处理中医“同病异治”和“异病同治”以及干预措施变异性大等问题,都有待于方法学上的进一步研究。而西医近年来强调的个性化治疗(personalized treatment)与中医的个体化治疗是有一定区别的。个性化治疗是指每一个个体自出生之日起就建立起一本个人疾病基因谱档案,针对其基因的易感性或危险性,决定针对某病所需采取的预防和(或)治疗措施,这些措施是靶点特异性的,只针对某一类人群,相当于亚组或亚群的治疗,因此与“同病异治”和“异病同治”概念不同。但两者也有共同之处,那就是治疗病症皆以患者为中心,而非“病”或“证”。总之,循证医学的方法可以用于中医证候诊断的客观化研究以及中药和中医疗法疗效、安全性和成本费用的评价,并有助于评价指标体系的建立。

利用客观证据以获得国际医学界的认同,能在更大范围内发展中医药学。循证医学的具体应用包括:(1)在中医药研究中开展国际认可、高质量的随机安慰剂对照试验,评价疗效的指标应以疾病终末事件发生为标准;(2)对现有高质量的临床试验加以收集和整理,根据不同的疗法和病种建立相应的疗效资料库,以促进国际交流与传播;(3)采用系统性评价的方法对以往发表的临床试验尤其是随机临床试验进行荟萃分析,为中医药走向世界提供确凿的科学证据;(4)正确认识和理解中医药疗法在防治疾病作用上的优势,对中医药研究疾病的优先性进行排序,把有限的卫生资源用于独具特色的疗法上;(5)对中医药治疗进行费用-效益的经济学评价,为医疗决策提供科学依据。

2 中医药临床试验的现状

中医药临床试验是 20 余年前才开始的,近 10 年来呈增长趋势。国内期刊文献中发表的中医药临床试验论文数量逐年增多,但在临床研究中仍存在不少问题^[1,2]。这些问题包括:研究设计与报告的质量不高,随机方法应用不当,没有足够的样本量,观察指标不明确;无论是证候或是疗效判断指标都难以达到规范化和量化;报告的疗效可重复性低,且疗效指标多为临床症状等“软”指标,缺乏长期随访所获得的终点“硬”指标(如病死率、致残率等)。这些问题均影响了研究结果的可靠性,因此其试验的科学价值很难得到国际认可。分析表明,中医药随机对照试验文章在设计、实施和报告中均存在一些问题,诸如随机分组方法的描述、盲法的使用、依从性及疗效定义等方面。

随机对照试验是目前国际上公认的评价干预措施效果的金标准,将其应用于中医药临床疗效评价具有重要的意义。我国《新药审批办法》规定,Ⅰ、Ⅱ期临床试验须采用随机方法。多中心、双盲、随机临床试验是国际临床试验研究的发展趋势。随机临床试验可用于评价两种干预措施的优劣,确定某一干预措施的利弊,证实某一干预措施的有效性和安全性。因此,严格设计的随机试验将对干预措施的效果做出肯定或否定的结论,通过推广应用有效的治疗和摒弃无效的治疗,能够起到节约医疗卫生资源、避免低水平重复研究造成的浪费以及提高医疗质量的作用。按病种或治疗方法进行系统评价,能够为医疗实践及正确地进行科研选题提供可靠的依据,并有助于确定临床相关结局的评价指标,为新药开发提供线索。

3 临床试验的随机化

中医药临床试验报告中存在“随机化”概念误用和滥用现象^[3]。随机对照临床试验对参与研究的对象进行随机化分组,其目的是为了人为的选择性偏倚,使进入试验的两组或多组在基线特征上尽可能保持一致,即统计学上的组间可比性,这样才能保证试验结果的真实可靠。临床试验中的随机化包含两层意思:一是指通过某种方法获得随机分配的序列,即随机分配方案的产生(generation of allocation sequence);二是指对该随机分配序列在实施分配期间进行隐藏(concealment of allocation sequence)^[4]。只对患者用随机方法进行分组是不完整的随机化,只有实施了分配方案的隐藏才称得上是完整的随机化。随机化方法学研究表明,未进行恰当随机隐藏的试验将可能夸大疗效 30% ~ 50%^[5]。

虽然以上随机分配方法可以避免偏倚的影响,但小样本试验(200 例以下)却很难保证各组患者在预后特征的分布上达到完全均衡。例如,一组患者中可能因为机遇(随机误差)的原因导致年龄分布的不同。一般来说,试验的样本量越小,出现预后分布不均的概率越大。由于存在机遇因素,因此组间的不均衡使试验结果的解释更加复杂。解决办法是进行分层随机(stratified randomization)^[6],即对一些影响预后和疗效的重要因素,如性别、年龄、病情等事先进行分层,然后再对每一亚组进行随机排列分配。多中心临床试验中的患者可按各中心进行分层。然而,分层的因素不宜过多,分层过多对试验样本量的要求将增大,同时也不利于结果的分析。

评价随机分配序列产生的标准:(1)恰当:试验报告描述了随机分配的方法,如计算机产生的随机数字、随机数字表或其他类似的方法。(2)不恰当:试验报告没有描述随机序列产生的具体方法,或描述的方法不适当,如采用按出生日期、住院日期或类似的交替分配方法。

随机隐藏的概念是指将随机分配方案对实施分配者在分组期间进行隐藏的手段。所谓恰当的随机隐藏是指中心随机化或由某一独立的单位进行随机,通过传真或电话告知随机中心办公室(国外多为临床试验中心)或医院药房的某一位独立的药剂师待分配患者,由他们通知实施分配者患者将接受何种治疗,从而使经管医生、护士或患者本人均无法选择所期望的治疗方案。中心随机方法是由计算机程序产生随机分配方案,然后由信息专家设计一个“按键式语音应答”随机方法,可以 24 h 在无人情况下通过电话与计算机之间的交互应答完成随机分配。我国现行最常用的方法是采用密封的、不透光的、按

序列编码的信封法,即参加临床试验的医生根据密封的信封内随机产生的治疗方案进行分配。由于临床医生可能一次同时打开几个信封或提前打开信封,结果是有的患者会被延期分配以达到接受期望治疗的目的。因此,这种隐藏并不理想,在发达国家已逐渐被中心计算机随机所取代。目前,在我国比较可行的方法是由中心药房控制随机分配,首先由接诊医生告知药房待分配的患者,再由药房负责试验的人员按预定方案进行随机分配。

如果由临床医生自己制作一套随机分配方案,并且由自己亲自参与患者的分配,则不属于随机隐藏,而被称为开放式试验。评价随机分配隐藏是否恰当有两个判断标准:(1)恰当:试验描述了分配隐藏,且采用的方法是适当的,即采用中心独立单位的随机或序列编码的、密封的、不透光的信封法。(2)不恰当:未描述分配隐藏或采用的方法不适当。如采用开放的随机数字表或交替分配患者的方法。

4 临床试验的对照、盲法和安慰剂

对照的设置决定了试验的类型。优越性试验(superiority trial 或 inferiority trial)通常采用安慰剂或不干预作为对照;而非劣性或等效性试验(non-inferiority trial 或 equivalence trial)通常采用阳性药物作为对照。美国食品与药品管理局规定,新药必须有与安慰剂对照的证据。

有关安慰剂对照的伦理学问题,不同的文化背景有着不同的观念,多年来在国际上一直存在着争议。2000 年修订的《赫尔辛基宣言》对此有如下描述:一项新的预防、诊断或治疗方法的获益、风险、负担和效果应当对比那些现行的最佳方法进行试验,但并不排除在缺乏有效证明的预防、诊断或治疗方法情况下使用安慰剂对照或空白对照^[7]。对于一些有生命威胁的疾病如癌症、严重感染等,则不主张患者接受安慰剂治疗,但如果所有患者均接受标准治疗,在此基础上比较试验药物与安慰剂则是被允许的,即 A + C 与 B + C 的比较模式(A 为试验药物, B 为安慰剂, C 为标准治疗)。临床试验中使用安慰剂伦理学的合理性必须考虑以下三个方面的因素,即试验的患者、已知现有治疗的效果以及该研究所要回答的科学及临床问题。在与标准治疗比较的临床试验中,如果试验药物与标准治疗之间没有统计学差异,并不意味着试验药物治疗该病就是有效的。因为,等效性临床试验所要求的样本量远比安慰剂对照试验要大得多^[8]。因此,如果缺乏足够的样本量(即统计学上的把握度),那么对等效性试验无统计学差异的结果就不能轻易下有效的结论。

一般来说,临床试验使用安慰剂应至少满足以下三个条件之一:(1)试验治疗的疾病属于自限性疾

病,即该疾病患者在不治疗的情况下大多会自行痊愈,如普通感冒、急性病毒性肝炎等;(2)该病症目前尚无特效的治疗方法,须探索新的治疗手段,如艾滋病病毒感染、某些病毒性或自身免疫性疾病等;(3)某些慢性病,其自然病程呈反复波动或变化,如高血压、糖尿病、哮喘、类风湿性关节炎、偏头痛、抑郁症等,短期不治疗不至于严重影响疾病的预后。

临床试验中盲法的使用被看作是评价试验质量的重要标准之一。有系统评价结果表明,采用双盲法与未采用双盲法的临床试验比较,后者会夸大治疗效果约 15%^[9]。盲法是指对参与试验的人员进行遮蔽,使其对实施干预(通常为药物)或评估结局的对象不知情。根据“盲”的程度不同可分为单盲(即患者不知道接受何种药物治疗)、双盲(患者和给予治疗的医务人员不知道接受或给予何种药物)、三盲(在双盲基础上对评估结局者设盲)、四盲(在三盲基础上对资料处理和统计分析者设盲)、五盲(在四盲基础上对撰写试验报告者设盲)。设盲的目的是为了避免实施偏倚(performance bias)和测量偏倚(detection bias)。但在某些情况下,使用双盲法是不可能(如外科手术或操作治疗,包括针灸、拔罐、按摩等)、不恰当(如针灸与药物治疗的比较、注射剂与片剂的比较)和不必要的(如评价的结局为死亡或存活)。对于前两种情况,若评价的结局带有主观判断(如评价症状改善情况),则应对结局评估者设盲,使其对患者接受何种治疗处于盲的状态。

5 临床试验的样本量

临床试验报告中有无预先的样本量估计是评价试验质量的重要依据之一。在试验设计阶段必须确定研究所需的病例数(通常称为样本含量)。理论上,验证某一干预措施与对照之间的差异,样本量越大则试验结果越接近于真值,其结果就越可靠。但由于资源的限制和伦理的原因,临床试验的对象数量不可能做到无限大,因此须确定符合统计学检验水准要求的最适样本量。试验时,须从研究对象(如患者群)中选择具有代表性的合适样本,将符合纳入标准的病例在经过其本人的知情同意后分组。临床试验的样本量过小,无论试验结果是否存在统计学差异,均不能排除因机遇(随机误差)因素造成的假阳性或假阴性错误。因此,小样本临床试验不能下肯定或否定的结论。目前,国内发表的临床试验大多为小样本试验。我们在对中医药随机临床试验进行评价中也发现,大多数试验的样本量都不大,且极少有报道预先计算样本量的试验^[10]。

试验设计时如何确定合适的样本量大小呢?一般可通过统计学方法估计样本量。根据试验目的和测量结局指标不同,计算方法亦不相同^[10]。优劣性

临床试验的目的是为了验证试验干预与对照干预效果之间是否存在差异,通常验证试验干预的效果优于对照干预。等效性临床试验的目的是为了验证试验干预与对照干预之间的效果相当,即差异无统计学意义,通常用于比较不同的有效治疗(如不同的抗生素)的疗效,也用于比较同一种药物不同剂型、不同给药途径的疗效。上述两种试验评价疗效的指标通常可以分为两类:(1)计数(定性)指标,如死亡与存活、阳性与阴性、正常与异常;(2)计量(定量)指标,如血压、血糖值、血清酶水平等实验室检测指标。有时,临床试验评价的结局指标有多个,估计样本量时需要选择其中最重要的结局指标。

6 中医药临床干预试验

中医药的临床干预大体分为药物与非药物类。中药可分为复方中药、中成药、单味药或中药提取的活性成分。试验中有关药物干预的描述应包括药物的组成、药材的来源、加工处理的方法、产品或制剂的质量控制、使用的剂量与疗程以及给药途径等。对于涉及知识产权的药物组成,可采用编码的形式,或只给药名而省略用量等方式解决。非药物干预如针灸治疗,应报道所采用的穴位名称与数目、针具的尺寸与材质、刺激的手法、留针情况、针灸师的资格和经验以及治疗方案和疗程等。同样,对于作为对照的干预措施也应如实描述。

7 临床试验的结局评价

临床试验的目的是为了验证干预措施的疗效和安全性。选择什么样的疗效指标(也称作结局或终点)是试验设计阶段需要考虑的重要问题^[11]。疗效的体现包含两层意义,即相对效应和绝对效应。相对效应是指试验干预与对照干预进行比较的相对效果;绝对效应是指干预措施本身在试验对象上所产生的有益或有害的净效应。两种治疗措施比较所产生的差异(效应)有三种可能的解释:(1)两种治疗本身存在真实差异;(2)机遇的作用,即由于临床试验的样本量较小而产生抽样误差所致,这种情况可以通过加大试验样本量来解决;(3)由于偏倚造成的假象,即由于试验设计和实施的质量较差以及未采用严格的随机分配和双盲法,致使观察结果偏离了真实的效应。因此,临床试验中干预措施效果的评价不仅要选择相对效应指标,同时也要选择绝对效应指标。这些结局测量指标可分为三种类型:定性指标(如临床痊愈)、定量指标(如住院时间)和时间序列指标(如生存率和复发率等)。最简单的定性指标属两分类资料,如有效与无效、好转与恶化、阳性与阴性、副作用有与无等。定性指标有时也可以为等级资料,如治疗恶性肿瘤的疗效分为:完全缓解、部

分缓解、无变化或恶化四级。定量指标比定性指标更加精确。如高血压的降压治疗,以舒张压降至 90 mm Hg 为有效,超过 90 mm Hg 为无效,那么 95 mm Hg 与 120 mm Hg 虽然均超过了 90 mm Hg,但却不能反映差异大小。因此,以实际测量值的定量指标表示效应的大小会更加准确。

所谓疾病的结局是指干预后健康状况的改变,如死亡与存活、痊愈与恶化等定性的终点结局;或者是连续变量的替代结局,如生化指标血清转氨酶、血脂、血糖水平等。急性病与慢性病在选择评定疗效的指标时有所不同。急性病常以痊愈率、病死率等作为评定疗效的指标;慢性病则常以缓解率、并发症发生率、生存率、复发率等作为评定指标。临床试验中比较两种干预措施的相对和绝对效益,过去常常采用卡方检验或 *t* 检验的统计学分析方法比较差异有无统计学意义。但近年来,国际上已开始采用以下一些效应指标计算疗效的差异和大小。选择何种指标需要根据资料的性质(定性或定量)和临床意义来决定。

7.1 相对效应指标及其意义 计数资料常用的指标有相对危险度(relative risk, RR)、比值比(odds ratio, OR)及 95% 的可信区间(confidence interval, CI)。在计算效应指标之前,应先明确两个概念,即试验组事件率(experimental event rate, EER)和对照组事件率(control event rate, CER)。EER 是指试验组治疗的患者中发生某一不良事件的病例占该组总病例数的百分比;而 CER 则是指对照组治疗的患者中发生某一不良事件的病例数占该组总病例数的百分比。如果以 *a* 代表试验组发生事件的病例数,*b* 代表没有发生某事件的病例数,则 $EER = a / (a + b)$;对照组中有(无)某一事件发生分别以 *c* 和 *d* 表示,则 $CER = c / (c + d)$ 。 $RR = EER / CER$ 。比值(odds)的概念是指试验组或对照组中发生某事件的病例数除以该组中未发生某事件的病例数。因此,试验组的比值为 a / b ;对照组的比值为 c / d 。OR 的概念为试验组的比值与对照组的比值之比,即 $(a / b) / (c / d) = (a \times d) / (b \times c)$ 。效应指标 RR 和 OR 反映干预措施与结局之间的相关性。CI 是衡量研究结果准确度(或不确定性)的一个指标,表示对某效应值估计的把握度,通常用 95% 表示 CI 的范围。95% CI 的含义为有 95% 的把握认为该效应值差异位于此区间范围内,即据此得出结论其错误可能性只有 5%。 $RR < 1$,表明试验干预效果优于对照干预;若 95% CI 的数值范围包含 1,则表明试验干预与对照干预比较的差异无统计学意义($P > 0.05$);当 95% CI 大于或小于 1 时,则表明效应值的差异具有统计学意义($P < 0.05$)。通常,临床试验的样本量越大,则效应值的可信区间范围越小、准确性越

高;反之,试验样本量越小,则可信区间范围就越大、准确性越低。

7.2 绝对效应指标及其意义 药物治疗的目的在于预防不良结局事件的发生,如死亡、卒中或心肌梗死等。绝对效应指标反映了基线危险度水平,具有更大的临床意义。常用的两个指标为:绝对危险度降低(absolute risk reduction, ARR)和需要治疗的病例数(number needed to treat, NNT)。ARR 是指对照组事件率与试验组事件率之差的绝对值, $ARR = |CER - EER|$ 。NNT 是 ARR 的倒数,即 $NNT = 1 / ARR$,即预防 1 例患者发生不良结局事件所需要治疗的病例数。因此,NNT 值越小越好。

7.3 其他效应指标 作为疗效评价指标之一的生活质量(quality of life, QOL),是指个体在自身文化和价值体系背景下,对目标、期望、标准以及所关心事物的生存状况的体验。已有许多 QOL 量表用于包括症状体征在内的评价。其中,SF-36 量表是目前国际上应用最广泛的一种生活质量量表,它由 36 个条目组成,包括生理功能、生理角色功能、疼痛、社会功能、心理健康、情绪角色功能、活力和总体健康等方面。中医药临床试验涉及症状体征评价的应当尽量采用国际上通用的 QOL 量表进行测量和评价。

8 临床试验的依从性与意向性治疗分析

临床试验的依从性包含两层意思:(1)临床试验在实施过程中按照试验设计方案执行的程度,即有无偏离原先的设计方案及偏离的程度;(2)纳入试验并被随机分组的病例是否接受干预措施以及治疗和随访的完整性。依从性是评价随机对照试验报告质量的重要指标之一。临床试验应当报告依从性的问题,包括入组病例接受治疗的完整性、试验期间退出或失访的病例数及其原因。任何临床试验都不可能保证所有试验对象均完全按照试验设计方案的要求完成治疗或随访,必然存在依从程度的差异。一般来说,门诊患者的依从性较差而住院患者的依从性较好;干预措施为疗程短、使用方便、副作用小的治疗依从性较好,疗程长、使用不便和副作用大的治疗依从性较差。临床试验在设计阶段应当考虑依从性的问题,并制定一些提高或确保依从性的措施,如向参加试验的患者详细介绍试验目的、治疗方案及随访的重要性,建立良好的医患关系,减少不必要的检查和辅助治疗,定期核查治疗实施情况等。但无论如何,退出与失访将是不可避免的。发表的试验中应明确报告符合纳入标准的病例数、实际参与随机分组的病例数、完成治疗和随访的病例数,并说明试验各个阶段病例离开试验的原因。患者退出临床试验的原因较多,如有的患者最初符合纳入标准,后因

证实诊断有误或患者因对试验知情后拒绝参加试验;有的病例在治疗期间由于试验药物的毒副反应很大或病情恶化,而改用对照药物或其他治疗;有的患者未按照试验要求接受全程治疗,如没有按规定的剂量、服用方法或疗程进行用药;有的患者因搬迁离开当地或治疗中途转院等原因而中断治疗;有的患者虽完成了治疗,但在随访期间失访或随访不全。上述原因在临床试验中较为常见,试验者应当如实报告,不得省略或隐瞒^[12]。

随机对照试验中的意向性治疗分析(intention-to-treat analysis),是指参与随机分组的对象,无论其是否接受该组的治疗,最终应纳入所分配的组中进行疗效的统计分析^[12]。该原则强调,只要是参与了随机分配的病例,就应当纳入最后的结果分析。因为随机分配的原则是确保组间的可比性,如果排除退出和失访病例,只对所谓“资料完整”者进行分析就会破坏组间的均衡性。观察性研究结果表明,退出试验病例的疗效普遍比完成治疗病例的疗效要差。因此,施行意向性治疗分析,其目的在于减少偏倚带来的影响,使结果更加真实。以往对于不依从的病例在结局分析中简单地予以排除,已被证明会夸大试验治疗的效果。如果接受试验治疗的患者因疗效不佳或出现毒副反应而退出试验,排除这些可能无效的病例就会造成偏倚,从而影响结果的真实性。

意向性治疗分析根据结局指标的资料类型不同而方法各异。对于结局为计数资料的试验,由于不能确定患者治疗的结局,通常将试验组和对照组退出和(或)失访的病例作为治疗失败处理。另一种处理方法是治疗组退出或失访的病例计为治疗失败(无效),而将对照组退出或失访的病例计为成功(有效)处理,这种分析方法称为“最差情况的演示分析(worst-case scenario analysis)”。试验者可选择其中一种或两种方法进行资料的意向性分析。应判断退出或失访对结果的统计学检验有无影响,如果纳入与排除退出或失访病例对结果有显著影响,则研究者在下结论的时候要十分谨慎。值得注意的是,有的临床试验实施不同阶段的随访,由于随访时间不同,失访的病例数就会不同。此时,作者应当分别报告不同随访阶段失访的病例数,并分别进行意向性治疗分析,将未获随访结局的病例按最后一次随访观察到的结局向前推移(carry forward),即将失访病例前一次随访的结果作为后一次随访的结局进行意向性治疗分析。

9 结 语

有专家预言,21 世纪的临床医学将进入循证医学时代。中医药的现代化必须借助循证医学的方法

和思路,将分析与综合、微观与宏观辨证统一起来,以中医理论为指导,确定科研的选题、方案设计和方法,把试验结果与中医理论要素结合起来进行评价。前瞻性临床研究将对中医证候客观化的研究,方剂适应病症、疗效和评价指标体系的系统研究,针灸的疗效和方法学评价等各方面均有重大的推动作用。

循证医学对中医药临床评价将在以下几个方面促进中医药的现代化进程:(1)促进中医药临床研究水平的不断提高,并最终与国际接轨;(2)制定循证的中医药临床评价指标体系,客观、科学地对中医药疗法的有效性和安全性做出评价;(3)通过在临床医疗实践中应用循证医学评价的证据,促进中医的临床实践,提高诊疗的效率;(4)借助相关的国际组织,开展中医药国际协作研究,学习并引进先进的研究方法和组织管理机制。

[参考文献]

- 1 刘建平. 中医药研究随机对照试验质量的现状及对策. 中国中西医结合杂志, 2003, 23(1): 62-64.
- 2 詹思远, 唐金陵, 谢立亚, 等. 中医药学术期刊随机对照临床试验文章评阅及建议. 中国中西医结合杂志, 1999, 19(9): 568.
- 3 Liu JP, Kjaergard LL, Gluud C. Misuse of randomization: a review of Chinese randomized trials of herbal medicines for chronic hepatitis B. Am J Chin Med, 2002, 30(1): 173-176.
- 4 刘建平. 临床试验随机化的概念与方法. 中国中西医结合杂志, 2003, 23(4): 244-246.
- 5 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. JAMA, 1995, 273(5): 408-412.
- 6 Jadad A. Randomized controlled trials. London: BMJ Publishing Group, 1998. 1-9.
- 7 World Medical Association. Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA, 2000, 284(23): 3043-3045.
- 8 刘建平. 中医药临床试验的对照与盲法设置. 中国中西医结合杂志, 2003, 23(10): 789-791.
- 9 Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. BMJ, 2001, 323(7303): 42-46.
- 10 刘建平. 临床试验样本含量的计算. 中国中西医结合杂志, 2003, 23(7): 536-538.
- 11 刘建平. 临床试验效应的表达——指标及其意义. 中国中西医结合杂志, 2004, 24(8): 677-679.
- 12 刘建平. 随机对照试验的依从性和意向性治疗分析. 中国中西医结合杂志, 2003, 23(12): 884-886.

[收稿日期] 2005-11-16