

## · 指南 ·

## 中医药与中西医结合临床研究方法指南

中国医师协会中西医结合医师分会 中国中西医结合学会循证医学专业委员会  
(2015 年 6 月 20 日)

## 1 背景

近年来,国家对中医药科研的投入力度逐渐加大,科研产出受到关注。国内中文期刊发表的大量临床研究报告质量较差,行业的认可度多数不高,能够真正进入指南推荐的甚少。其原因之一在于缺乏临床研究方法学指导。目前国内尚未见系统、全面的指导中医药和中西医结合临床研究的方法学指南。这也是本指南制定的初衷和当前行业的需求。

本指南旨在帮助中医药及中西医结合临床科研人员进行临床研究方法学设计和实施,重点介绍中医药及中西医结合领域常用的临床研究方法,包括随机对照试验(randomized controlled trial, RCT)、队列研究(cohort study)、病例对照研究(case-control study)、病例系列研究(case series)以及与临床研究相关的其他研究方法等。该指南所涉及的范围包括中医药各类疗法,如辨证论治、复方或单味中成药治疗、针灸疗法及太极拳和推拿等非药物治疗方法,以及中医药各种疗法与西医药疗法不同方式的联合应用,开展有效性、安全性及卫生经济学的评价。本指南也适用于其他传统医学包括民族医学及民间疗法等临床研究。

## 2 目的

根据中医药及中西医结合整体观综合干预、个体化诊疗的特点,介绍 RCT、队列研究、病例对照研究、病例系列研究的设计、实施要点及应用范围;提出使传统中医学体系和现代试验研究工具结合的方法和建设,从而获得合理的临床研究证据。本指南力求规范中医药临床研究,为中医药和中西医结合临床科研提供方法学指导,产出高质量的循证证据,可供临床医疗决策和政府卫生政策的制定作为参考。该指南也可以作为临床科研人员培训的参考教材

通讯作者:刘建平, Tel: 010-64287002, E-mail: jianping\_l@hotmail.com; 陈可冀, Tel: 010-62860894, E-mail: kjchenvip@163.com

DOI: 10. 7661/CJIM. 2015. 08. 0901

使用。

## 3 中医药与中西医结合临床实践的特点

中医药理论体系在整体观的指导下,以辨证论治为诊疗特点。中医临床实践中的常用疗法包括药物疗法和非药物疗法。中西医结合则是将传统的中医药知识和方法与现代医学知识和方法结合起来,力求优势互补、提高临床疗效,并在此基础上阐明其作用机理,达到传承、创新、发展中医药学的目的。中医药与中西医结合临床研究在很大程度上采用现代科学知识和技术对中医药诊疗过程中临床实践作出科学的阐释,同时兼顾中医药体系的特点,探索在共性规律基础上的个体化医疗效果。

## 3.1 遵循中医药传统理论的思维和原则

中医药学理论是指导中医药临床实践的重要理论依据。中医药临床研究中,应遵循中医药传统理论的思维和原则,阐述所选择的治疗方法的合理性,这也是中药与植物药临床研究的区别点。在中医药临床研究中,特别是中药复方、单味药物及其提取物或活性成分的临床应用研究,力求符合中医药传统理论的思维和原则,体现其临床应用特点。中西医结合临床研究应重视综合干预(复杂干预)的研究特点,进行方法学设计。

针灸临床研究也要遵循中医药传统理论,体现中医针灸临床实践的特点,探索循经辨证治疗、不同针法及不同灸法治疗等在临床治疗中的效果。其他非药物疗法包括八段锦、太极拳、推拿、气功等临床研究,也要充分考虑中医药理论特点和合理的方法学设计。

## 3.2 辨证论治原则下的中医药临床研究模式

辨证论治是中医药治疗学的精华所在。在辨证论治的原则指导下,中医药的临床研究常采用以下几种模式:病证结合的临床研究、“病”的临床研究及“证”的临床研究。

**3.2.1 病证结合的临床研究模式** 在疾病共性规律与患者个体特征有机结合的基础上进行临床研究,是目前开展最多的中医药临床研究模式。病证结合临床研究中的“病”是指现代医学的疾病,“证”是指中医的证候,其临床疗效评价也应采用病证结合的观察模式。既要选择行业公认的疾病疗效评价指标,评价干预措施对“病”的疗效,也要采用主客观结合的评价方法评定中医证候的疗效,目前,多应用症状积分量表的改善或主要症状的消失率等进行观察。鼓励选择能够体现中医学治疗特色与疗效优势的评价指标,尽可能将“望、闻、问、切”信息的量化与中医辨证指标结合,建立共识度较高的中医证候疗效评价的方法。所有中医药与中西医结合疗法均可开展病证结合的临床研究。

**3.2.2 “病”的临床研究模式** 强调评价干预措施对现代医学疾病的有效性和安全性。中医辨证可以作为入组的诊断标准,和(或)指导制定治则治法的标准,或作为疗效评价中亚组分层的依据,但中医证候则可不作为主要的疗效评价指标,而强调治疗方法对现代医学疾病的疗效。疗效评价指标应选择临床终点结局指标、患者相关结局指标如生活质量、患者报告的结局(patient-reported outcome, PRO)等以及现代医学公认的疾病相关替代指标。所有

中医药与中西医结合疗法均可开展针对现代医学疾病的临床研究。

**3.2.3 “证”的临床研究模式** 在中医辨证论治原则的指导下,评价中医证候类药物的有效性及安全性。应重视中医证候的标准化和客观化,要选择公认的证候评价标准或量表,重视证候疗效临床价值的公认性,应充分说明其科学性。选择中医证候作为疗效评价指标的临床试验,通常要建立证候积分定量评价的工具或量表和质量控制体系。推荐临床研究开始前进行不同研究者和不同研究中心之间证候定量评价的一致性检验,以确保评价结果的可信性和一致性。可开展中药复方以及针灸等非药物疗法对中医证候影响的临床研究。

此外,还可以开展针对构成中医证候主要“症状”的临床研究模式。

### 3.3 综合干预模式下的中医药临床研究特点

临床实践中,中医药疗法往往只是作为综合干预(复杂干预)的一部分,这种综合干预通常包括标准化西医的联合治疗以及饮食和生活方式干预等,这些干预可以同时实施,也可以先后实施。临床观察到的疗效往往是综合疗效的结果,因而在综合干预(复杂干预)模式下的临床研究中,得出的“疗效”不仅是中医药的疗效,还要分析多种因素对疗效的影响。在综合干预模式下,临床研究的结果可能会由于不同疗法之间同期应用而产生以下相互作用:疗效增加(协同及相加作用)、疗效降低(拮抗作用)、毒副作用增加、毒副作用降低等。要根据干预措施的特点、前期的研究结果来确定研究目的,说明评价综合干预设计的合理性。

在综合干预临床研究中,要根据研究目的选择不同的设计方法,如要评价中西药联合治疗的整体疗效,且西药的使用符合相应的指南或规范,可采用加载设计(对照组在西药基础上设立中药的安慰剂组),此类设计可采用随机双盲对照试验。如安慰剂无法设置,则可选择实用型随机对照试验设计(pragmatic randomized controlled trial, pRCT),即对照组仅为西医治疗。在随机分组无法实施的情况下,如患者对治疗方案有明确的选择偏好,则可采用前瞻性队列研究设计。在此类临床研究中,除了应用常规统计分析方法之外,应当采用多因素分析探讨多种影响疗效的因素。

## 4 中医药及中西医结合临床研究设计的基本原则

### 4.1 伦理与受试者保护

尊重、保护受试者的权益、安全和健康是临床试验伦理学的基本原则。进行临床试验必须有充分的科学依据。临床试验的设计与实施应符合《赫尔辛基宣言》等国际公认的伦理原则,同时要符合我国《药物临床试验质量管理规范(good clinical practice, GCP)》、伦理委员会具体工作细则和审查技术原则,可参照国家食品药品监督管理局(State Food and Drug Administration, SFDA)颁布的《药物临床试验伦理审查工作指导原则》等相关伦理要求。对受试者的权益、安全和健康的考虑必须高于对科学和社会利益的考虑。

### 4.2 对比

设立对照是科学的核心思想和基本要求,只有通过比较才能有鉴别。在临床疗效研究中,接受治疗的患者病情好转不等于治疗措施一定有效,影响疾病转归的因素有疾病的自然史、患者的体质和营养状态、非特异性因素、治疗措施等。因此,应根据临床试验的需要设立相应的对照组包括空白、安慰剂、标准治疗或常规治疗等。

### 4.3 均衡

为了保证对比研究结果的准确性、可靠性,排除受试对象组间

差异所带来的影响,合理地推断因果,有必要采取措施如随机分组或配对确保受试对象在组间基线特征上均衡可比,也就是确保试验组和对照组患者的背景相同或近似。

### 4.4 重复

单个病例或少量病例的观察结果容易受到机遇的作用产生随机误差。因此,临床研究如经过多次重复可以获得相同或相似的结果,则其研究结果具有可重复性。一般来说,重复的次数越多,即样本量越大,机遇对结果造成的影响越小,从而使结果更接近客观真实情况,但并不是样本量越大越好。因此,临床研究需要预先估算样本含量。

### 4.5 盲法

盲法是临床研究中十分重要的设计原则和质量控制措施。根据盲的程度不同,可以使研究对象、研究者、结局评价者等研究相关人员不能预先知晓其研究对象所接受的干预措施,以避免信息采集、结局测量的过程产生偏倚,保证测量数据的真实性,减少测量过程中主观因素的干扰。

## 5 中医药与中西医结合临床研究设计要点

中医药与中西医结合临床研究方案的设计,一般需要考虑以下内容:立题依据、明确临床研究目的、制定临床研究方案、确定适宜的目标受试人群、明确干预措施及其对照、选择有效和安全的评价指标、建立临床研究质量控制措施等。

### 5.1 立题依据充分

临床研究的首要问题是立题是否充分,应说明研究问题的临床价值。在立题之初,要对既往的研究进行整理、归纳和系统综述,找出临床研究的切入点和创新性,说明其优势和进行临床研究的科学价值。

### 5.2 研究目的明确

确定一项临床研究目的需依据整体临床研究计划考虑。首先应明确该项临床研究所处的临床研究阶段,是探索性的研究,还是验证性研究,不同阶段的临床研究其目的不同。其次,应充分评估既往临床前研究或临床研究所提供的数据,明确将要开展的下一项临床研究所需回答的问题,以保证各项临床研究之间合理有序地衔接。试验目的是设计和制定临床研究方案的前提。

中医药和中西医结合临床研究目的应具体,具有可行性,符合临床实际,并突出中医药临床优势与特点。确定研究目的要依据干预措施的功能主治和作用特点,参照既往临床研究工作基础,体现干预措施的临床应用价值。干预措施功能主治的表达要根据干预措施临床作用特点制定,一般分为三种情况:一是以中医的证候作为临床主治;二是以现代医学的疾病和中医证候结合作为临床主治;三是仅以现代医学疾病作为临床主治。目前中医药的临床研究功能主治范围多以病证结合为主体,要结合临床实际确定中医药的功能主治,以此作为临床研究所遵循的主线。

### 5.3 临床研究方案设计(表1)

在中医药的临床研究中,要注意结合中医药的特点和需要回答的临床研究问题,适用于西药的 RCT 模式不能完全简单地照搬到中医药临床研究中。建立科学假说、确定研究目的是制定研究方案的前提。针对中医综合干预(复杂干预)的评价可参考英国医学研究委员会评价复杂干预措施的方法指南。建议采用分阶段的研究方法:(1)回答中医药治疗特定疾病具体效力的有关问题可以参考解释性随机对照设计(explanatory randomized controlled trial, eRCT)。这类研究的特征是高度选择的同质性人群,标准化的干预措施,以安慰剂作为对照,采用双盲或双盲双模拟设计的试验;

(2) 回答有关中医药干预在常规临床实践中效果的问题可以参考 pRCT; (3) 在前期研究阶段回答有关中医药的疗效, 可考虑包含个体化治疗的研究, 如观察性研究、病例系列研究(预试验); (4) 如果前期研究的数据显示疗效可观, 后续的研究可采用标准化的治疗方案, 如采用固定处方治疗疾病; (5) 其他问题的解决可能不需要 RCT, 可通过其他研究方法进行探讨。

表 1 研究问题及其相应的研究方法

研究问题	优选的研究方法
中医药治疗某种疾病或症状的效力	安慰剂对照, 双盲, RCT
中医药在现实世界中的效果	pRCT
不良反应是否由中医药导致	观察性研究-纵向调查, 病例对照研究
患者服用中医药的经历	定性研究
中医药疗法对哪些症状效果良好	受试者注册, 观察性研究, 横断面调查
特定中药的有效成分	实验室研究
针灸干预起效机制	实验室研究
中医药治疗疾病的经济效益	卫生经济学评价

#### 5.4 确定受试者

受试者的选择根据临床试验目的决定。选择受试者是实施中医药临床研究的重要环节。对于病证结合模式下的临床研究, 建立正确的疾病与证候的诊断标准是确保样本同质性的关键, 尤其是多中心临床试验, 试验设计中应确定统一的目标适应证诊断标准(包括现代医学疾病与中医证候)、纳入标准、排除标准。其中, 诊断标准应为能够正确诊断一个疾病或证候的现行公认的标准, 包括西医诊断标准或(和)中医辨证标准等。纳入和排除标准是确定选择可能获得研究结果的受试者, 排除可能混淆结果的受试者, 并排除可能处于高风险的受试者(伦理学的考虑)。

#### 5.5 样本量估算

样本量的估计是临床试验设计的关键点之一, 每个临床试验所需的样本量最低应满足统计学的要求, 以确保对试验目的进行可靠回答。临床研究早期阶段的探索性研究可以不进行样本量计算。样本的大小通常依据试验的主要结局指标[疗效和(或)安全性终点]来确定, 同时应考虑试验设计类型、比较类型、检验假设、I类和II类错误参数等。

#### 5.6 确定干预方案

临床研究方案中需制定干预措施, 包括干预方案、剂量、实施途径、疗程、合并治疗的规定等。中药临床试验的干预剂量通常根据既往临床应用经验或前期研究结果确定。安全性也是干预剂量设计时需考虑的重要因素。疗程应根据疾病的发展变化规律和干预措施的作用特点确定。合并治疗包括基础治疗和联合治疗。合并治疗应预先规定, 否则会严重干扰有效性和安全性的评价。

#### 5.7 选择合理的对照

对照的设置需要根据研究目的确定。并非所有的临床研究都需要设置对照组, 如无对照的观察性研究, 在自然状态下对研究对象特征进行观察、记录, 并对结果进行描述和分析。但在疗效评价的比较性试验研究中, 需设计对照组来比较研究因素对研究对象造成的影响。在等效性试验中, 中医药治疗可与标准西医治疗或其他中医药治疗等方法对比。在剂量-效应研究中采用不同剂量互为对照。当研究中医药的确切疗效时, 在伦理学允许的情况下, 可选用安慰剂作为对照, 以获取中医药治疗效力的基线测量资料, 基于此再通过后续比较评估该阳性干预措施是否优于或不劣于其他治疗措施。安慰剂对照的选择应该谨慎, 选择的标准取决于整个研

究的设计是 pRCT 还是 eRCT。前者的目的是为了提供更加贴近临床实际情况的综合性治疗效果(effectiveness), 需要与当前的最佳常规治疗相比较, 便于临床医生的决策。后者的目的是为了评价除安慰剂效应以外与干预措施相关的特异性疗效(efficacy)。如果安慰剂的选择和设置不合理, 研究将受到质疑。

#### 5.8 确定结局与指标

效应观测与评价指标包括有效性评价指标和安全性评价指标。有效性指标是反映干预措施作用于受试对象所表现出的有效性的主要观测与评价工具, 主要包括疗效观测指标及其判定标准。病证结合模式下的中药有效性评价, 主要包括疾病有效性评价和中医证候改善的评价。可以选择公认的临床终点指标、患者相关的结局指标或替代指标, 也可以根据试验目的选择其他适宜的指标。针对中医证候疗效的评价, 临床试验中尽量采用经科学研究、信度和效度检验的中医症状量表。

安全性指标观测与评价要根据研究药物或疗法的目标适应证、纳入受试人群的特点、疗程、干预途径、已知毒性靶器官和既往临床应用经验等全面设计, 并有足够的暴露时间及病例数以评价其安全性。根据临床试验目的, 安全性指标也可以作为主要效应指标。

#### 5.9 随访的考虑

中医药在远期疗效和健康促进等方面可能具有一定的优势, 可考虑设置足够长的随访时间, 以便更客观地评价中医药的疗效。根据药物或中医疗法不同作用特点和试验目的, 设计随访的期限与次数、间隔时间, 随访可以针对进入试验的所有受试者。如果随访目的是观察疗效的稳定性及疾病复发情况, 可以只选择观察临床疗效为痊愈和(或)显效的病例数据。

#### 6 中医药与中西医结合临床研究质量控制要点

良好的临床研究质量是保障一个设计恰当的临床试验获得可进行评价的有效性和安全性数据的必要条件。中医药临床试验在符合 GCP 原则的基础上, 在试验设计时应预先考虑到影响数据采集的关键点, 在临床试验实施前制定质控措施和标准化操作流程(standard operation procedure, SOP), 这均需要在临床试验方案或临床试验的相关文件中提前规定和说明。重大临床研究项目应当建立相应的数据安全监察委员会(Data and Safety Monitoring Boards, DSMB)和数据协调中心(Data Coordinating Center, DCC)。

在实施过程中应加强临床试验质量控制。现阶段要重视控制以下影响中医药临床试验质量的主要因素。

##### 6.1 对受试者选择偏倚的质量控制

当某一研究结果与其真实值之间出现了某种差异, 这种差异的现象及结果称为偏倚。选择性偏倚产生在研究的设计阶段, 由选择研究对象的方法不正确所造成, 是干扰评估临床研究结果偏倚的重要因素之一。严格的科研设计强调采用随机化原则进行分组, 能有效地防止选择性偏倚。对选择性偏倚的质量控制要关注随机序列的产生和随机隐藏。要关注研究设计是否选择了合适的随机分配序列产生的方法, 试验过程是否真正实现了随机隐藏, 是否按照方案执行了既定的随机分配, 是否有破坏随机的情况发生等。研究者应当遵循试验方案的随机化程序, 保证随机化方法严格执行。

##### 6.2 对受试者评价偏倚的质量控制

在中医药的临床研究中, 由于中医药自身的特点, 疗效评价主观指标较多, 更应该重视通过合理设计控制评价性偏倚。盲法的设计和实施是控制评价性偏倚的主要措施, 临床试验根据设盲的程度分为开放(非盲)、单盲、双盲。双盲是指对受试者方(包括受试者本人、

家属、监护人、陪护人员等)和研究者方(包括信息采集人员、效应指标评价人员等与研究有关的临床研究团队的所有人员)都未知受试者分组情况,除此之外还包括监察员、数据管理员、统计分析人员等。双盲试验需要试验中所采用的处理方法在用药前或用药时都无法从感官上识别,且在整个试验中都保持盲态。双盲临床试验的双盲原则应自始至终地贯彻于整个试验之中。从方案制定、盲法设计中随机分配序列号所对应的治疗措施(盲底)的产生、受试者入组用药、研究者记录试验结果并做出评价、监察员的检查、数据管理直至统计分析,都必须保持盲态。任何非规定情况所致的盲底泄露,都会导致破盲。由于中医药自身的特点,导致安慰剂制作存在实际困难,应加强临床试验安慰剂的研究,避免因安慰剂制作质量导致临床试验实施中破盲。当试验药物和对照药的剂型、用法、用量不同时,则采用模拟技术,如双盲双模拟技术,即为试验药与对照药各准备一种安慰剂,以达到试验组与对照组在用药的外观与给药方法上一致。双盲试验需要保留盲法操作过程文件的记录,并应在临床试验总结报告中说明。

如果基于伦理学和可行性的考虑,不适宜采用双盲,则应考虑单盲试验或开放性试验。采用单盲或开放性试验均应制定相应的控制偏倚的措施,使已知的偏倚来源达到最小。如主要变量应尽可能客观,采用中央随机化方法以及采用盲法评定等。

### 6.3 量表或主观症状评价的质量控制

在中医药临床试验中,采用与疾病相关的症状、体征或症状群量表是有效性研究的重要部分。常见的问题是参加临床试验的研究者评价的一致性差,尤其作为主要疗效指标时,影响了对有效性的最终评价。在使用症状、体征或症状群量表评价有效性时,要重视对研究者使用评价一致性的质量控制。在临床试验方案设计时,应采用公认的、信度、效度和反应度良好的量表或症状等的量化标准。研究者在临床试验实施前应该接受统一的培训,培训后接受量表使用的一致性检验,如有必要,应用某些精神适应证领域量表,还应取得相应的资质,以保证研究者对量表使用的质量。尤其是多中心试验时,应确保试验前研究者针对量表及主观症状所采集的临床试验数据具有可评价性。

### 6.4 理化检测的质量控制

参与临床试验的医疗机构的临床检验实验室应当建立质量管理标准和 SOP,保证检测、诊断数据及结果的准确、可靠。要关注主要诊断指标或疗效指标多中心实验室的一致性和质量控制。推荐使用通过卫生部临床检验中心的室间质量控制评价的实验室和检测项目,或通过 ISO15189 认证的实验室。

如实验室检测指标为有效性或安全性评价的关键指标,或特殊的检查指标及变异性较大的指标,如病毒、结核菌、器官纤维化指标等,不同中心的实验室检查结果受检测仪器、检测条件及检查人员的影响,建议采用中心实验室一个批次集中检测。中心实验室可以选择专业机构认可的实验中心。选择中心实验室需要制定严格的样本采集、储存、配送的 SOP,保证样本采集和运送的质量及安全性。

若未使用中心实验室,由于设备、分析者和参考值范围的不同,不同的研究中心可能获得不同的实验室评估结果,此时需要采取措施取得一致性的数值,如进行检验方法统一培训和一致性测定,这对于以实验室指标作为主要指标时尤为重要。否则,有可能由于系统误差(偏倚)导致不同研究中心的实验室检查结果难以互相比较或合并,从而难以对研究药物的安全性和有效性进行正确评价。

对于非实验室检查指标,如医疗器械检查(如血压检查、心电图运动负荷平板试验等)或影像学检查(如 X 线片、超声检查、CT、MRI 等)的指标,要选择临床公认、质量可控的测量方法和测量仪器,对检查过程要制定规范的 SOP,保证不同中心、不同人员检查测量的一致性,检查过程中要控制影响检查结果的人为或仪器因素。明确对测定仪器的要求,明确对被测量人员和环境的要求,明确对测量人员的技术要求和培训规定等,并尽可能在临床试验设计之初制定相应的规范,如原发性骨质疏松症,应预先规定骨密度仪精确性误差测量规范。如检查结果作为主要疗效指标或疗效评价的关键内容,建议采取第三方统一进行。

### 6.5 临床研究数据管理的质量控制

临床数据管理是临床试验实施过程的重要组成部分,是数据收集、录入、清理和管理过程的综合体现,其目的是获得高质量的临床数据用于统计分析,从而得到真实可信的结论。整个数据管理过程包含了试验前准备阶段的制定数据管理计划、设计数据采集工具[研究病历/病例报告表(case report form, CRF)]、建立数据库、制订数据核查计划、培训临床研究人員等;试验中的数据采集、源数据现场核查、数据编码、数据清理;试验结束后的盲态审核、数据库闭合、数据管理文件归档等步骤。对临床研究数据的质量评价国际上也有明确的要求,美国 FDA 发布的临床试验中应用计算机系统技术指导原则(computerized systems used in clinical trials, CSUCT)对数据质量的要素作了明确的定义,包括归属性(attributable)、可读性(legible)、实时性(contemporaneous)、原始性(original)和准确性(accurate),合称为 ALCOA 原则。除此之外,良好的数据质量还应达到留存的持久性(enduring)、采集的完整性(complete)和内外的一致性(consistent),以上三点与 ALCOA 原则合称为 ALCOA<sup>+</sup> 原则,数据质量的高低取决于数据对于这些属性的展现程度。ALCOA<sup>+</sup> 是国际上评估临床研究数据质量的重要原则。我国 SFDA 基于以上原则,颁布了临床试验数据管理工作技术指南(2012 年),可以作为临床研究数据管理过程的指导和质量评价的参考。

### 6.6 临床原始数据采集的质量控制

通过临床研究原始数据的完整采集记录,可以了解影响临床研究质量控制的相关因素,也可以解释临床研究数据中发现问题,可以体现研究实施过程中是否依从方案,是否及时、准确、完整、规范、真实地采集了临床研究数据。中医药临床研究需关注如下原始数据的采集。(1)基线数据的获得。临床研究的基线筛选方法可以保证纳入符合要求的受试者,但应保留相关筛选记录,如纳入标准要求以初次诊断经过生活方式干预后的疾病人群,则受试者应保留生活方式干预的记录以及筛选的相关检查记录。(2)纳入受试者人群的证据。根据临床研究设计预先拟定的纳入受试者人群,在临床研究中应保留证明所纳入疾病人群特征的原始记录(如病史、病程、既往用药情况),重要的疾病诊断依据(如 CT、超声、心电图、冠脉造影、病理报告等),并通过恰当的原始 CRF,提醒研究者收集和整理重要的纳入疾病人群信息,以确保所纳入疾病人群符合设计的要求。(3)缺失数据的追踪。临床研究的数据常见缺失原因:受试者拒绝继续参与研究、治疗无效或有效、不良事件的发生、受试者出差等,部分病例可能仅有纳入时基线的测量数据,或一次或多次随访后数据缺失。虽然对于全分析集一般要求对未记录的数据进行截转,但实际上缺失数据是临床研究产生偏倚的重要原因,过多的缺失数据可能影响有效性的评价、影响试验组与对照组之间的可比性、影响所纳入疾病人群特征的代表性。因此,临床研究应追踪和

记录所有临床试验数据的信息, 尽量避免缺失数据。(4) 合并治疗的记录。在临床试验实施过程中, 为保证受试者的安全, 在符合伦理学原则前提下, 虽然预先规定合并治疗方案, 但临床试验期间难以避免使用影响有效性和安全性的药物, 应如实详细记录, 以评估其对有效性及安全性的影响。应注意的是, 过多影响有效性和安全性药物的合并应用, 将导致临床试验有效性和安全性结果难以评价。

### 6.7 多中心临床试验的质量控制

多中心临床研究可以在较短时间内收集研究所需的受试者, 且收集的受试者范围广, 用药的临床条件广泛, 研究的结果对将来的应用更具代表性。

在进行多中心临床研究时, 通常按照便利性和可获得性来选取研究中心, 对于选取具有目标受试者人群代表性的样本, 研究中心的选择很重要。在进行固定样本的多中心临床试验时, 需要关注在相同的样本量下, 选择较多的研究中心相对于较少的研究中心而言, 观察到显著的治疗-中心交互作用的概率更大, 从而影响疗效的判定。

多中心临床研究要求各中心研究人员采用相同的试验方法, 试验前人员应统一培训, 试验过程要进行质量控制。当主要指标易受主观影响时, 需进行一致性检验。若未使用中心实验室, 则由于设备、分析者和正常值范围的不同, 不同的研究中心可能获得不同的实验室评估结果, 此时需要采取措施取得一致性的数值, 如进行检验方法和步骤的统一培训和一致性测定。这对于实验室指标作为主要指标时尤为重要。否则, 由于可能产生的混杂偏倚导致不同研究中心的实验室检查结果难以合并, 从而难以对研究药物的安全性和有效性进行无偏估计。

### 6.8 DSMB

DSMB 是监督和指导研究者从事临床研究的独立的专家顾问小组。其成员来自于不同学科背景的专家组成, 包括医学(相关专业背景的医生)、临床药理学和(或)毒理学、流行病学、统计学、临床试验管理与伦理学专家等。其职责是为确保研究的科学性、完整性、数据的可靠性以及受试者的安全。

考虑建立 DSMB 的必要性时, 需要结合研究的规模、复杂度以及所研究疾病潜在的安全性风险。通常, 当研究满足以下几个特征时, 建立 DSMB 是十分必要的: (1) 大型、多中心、双盲的 III 期临床研究, 研究干预对受试者存在潜在的安全性风险; (2) 对敏感或弱势人群的研究, 或有风险的干预措施, 或在试验过程中可能出现高发病率或病死率的终点(end-points); (3) 当临床研究在紧急情况下实施, 而无法征得患者同意(patient consent); (4) 关注症状解除的大型研究; (5) 试验过程中可能存在较高的毒性风险或其他重要的医学风险; (6) 转基因试验。

DSMB 通过定期召开审查会议控制研究质量。审查内容包括: (1) 阶段的效果和安全数据; (2) 研究进展报告(各分中心的进展); (3) 严重不良事件报告; (4) 累积的安全数据; (5) 数据的质量、完整性、及时性; (6) 累积入组患者揭盲后的随机分布情况、人口学特征; (7) 方案的依从性(包括临床研究者及患者两方面); (8) 影响研究结局和危害数据、参与者机密的因素。

## 7 常用临床研究的设计方法

### 7.1 RCT

#### 7.1.1 定义

RCT 是在人群中进行的前瞻性的、用于评估医学干预措施效果的临床研究方法。按照正规随机方法, 使每位研究对象有同等的

机会被分入试验组和对照组, 试验组实施治疗措施, 对照组实施对照措施或仅给予安慰剂, 在相同条件下, 应用客观效应指标, 经一段时间随访观察后, 比较两组的差别。

#### 7.1.2 研究设计和方法学要点(图 1)

在实际的中医及中西医结合的临床研究中有多种 RCT 设计可供选择。

(1) eRCT。其设计在于能够控制所有可能存在的混杂因素的影响, 从而精确测评与安慰剂或阳性对照相比其干预措施的特定疗效。这种方法的优点在于可用于研究某种具体干预措施与效力之间的因果关系; 缺点则是可能不适用于评估受特定环境因素影响的复杂个体化治疗的疗效。eRCT 也常被用于检测在理想条件下, 某种单一疗法对经过精选的同类人群的疗效。这种情况通常不能在“现实世界”的临床实践中推广。

(2) pRCT。与 eRCT 相比, pRCT 对各种因素的控制相对宽松, 并且不试图排除治疗的背景效应。这种设计强调在现实世界中, 对异质性较高的人群采用以患者为中心的结局指标来检测某种疗法的实际效果。它的内部设计严谨度较低, 因此无法建立某种特定疗法与效力的因果关系, 但是 pRCT 比 eRCT 具有更高的外部真实性和外推性。

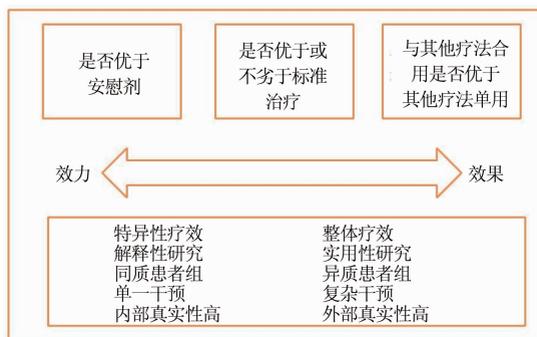


图 1 研究问题及其相应研究设计的关系<sup>[2]</sup>

(3) 交叉试验(cross-over trial)。是指在试验的第一个阶段, 按照随机的方法将患者分配到治疗组或安慰剂组接受相应的治疗, 之后经过一段无治疗措施的洗脱期, 然后接受与之前相反的治疗。每位受试者都会接受两种治疗, 只是接受这两种治疗的先后顺序是随机决定的。交叉试验的优点是将差异最小化。每位受试者都是自身对照, 因此个体间差异就可以最小化, 从而使需要的样本量更小。交叉试验只适合于研究病情稳定的疾病或者干预措施的短期疗效。尽管已经有一些非常成功的使用交叉设计来研究中药治疗特异性湿疹的试验, 但由于中药可能存在长期疗效, 这种试验设计可能并不适用于中药研究。

(4) 单病例随机对照试验(N-of-one RCT)。可以被认为是一名受试者的交叉试验。N-of-one RCT 有很多实施方法, 常用的方法是使受试者按照随机分组的方法接受 A 治疗或 B 治疗, 一段时间之后采用另一种可能有效的疗法或安慰剂。这种方法可反复多次进行从而来确证某种特定疗法的疗效。这种方法比使用传统病历的方法更加严格也更为科学。N-of-one RCT 可用于检测个体而非群体对于某种治疗的反应, 可调查基于药物选择和剂量改变的治疗变化。但是, 试验的结果并不能推论到其他患者而且一个阶段干预治疗的效果可能被带入下一个阶段, 这有可能是非常重要的混杂因素。

(5) 加载设计 (add-on design)。加载设计是中西医结合临床研究的一种设计方法,它是在现有临床标准治疗(如西医)基础上加用中药或安慰剂。一般在所研究的疾病已经有一种标准治疗并且被证实疗效,从伦理学考虑不宜中断原来的标准治疗时,考虑加载设计。

在采用加载设计时,所选择的标准治疗应被公认,疗效指标要明确和恰当,应能反映出加载的试验药物的添加疗效。受试者选择应有可比性,一般筛选出既往使用标准治疗已经取得最大疗效,但未达到治疗目标,同时病情保持稳定的疾病人群作为受试者。在采用加载设计时,应注意治疗的标准化和一致性,其中包括规定允许标准治疗的条件,允许使用药物的种类,使用的剂量、方法、时间等。观测指标选择应全面,除了评价受试者疾病主要疗效指标外,标准治疗药物的耗用量或使用频率及某些标准治疗已知不良反应发生的频率或严重程度改变,均可作为评价试验药物作用的指标。

由于加载设计得到的疗效是多种施加因素的综合结果,中药效应的确认易受混杂偏倚的影响。当出现罕见或不常见的不良反应时,往往无法确定是由哪种药物或两种药物共同造成的,受试者需要承担两种药物未知的混合作用的风险,解释有时显得较为复杂或困难。当标准治疗本身的疗效过高时,“天花板”效应可导致无法鉴别试验药物的疗效。基于此,应用加载设计需慎重,并仔细考虑标准治疗的确定和一致性、受试人群的选择等。

(6) 剂量-效应研究设计。剂量-效应研究是评价中药疗效,选择最佳治疗剂量的一种有效研究方法,其设计类型一般可以分为平行量效研究、交叉量效研究、强制剂量滴定和供选择的剂量滴定等。

平行量效研究是剂量研究中的常用设计方法。即将受试者随机分为多个有各自固定剂量的组。固定剂量指最终或维持剂量;受试者可开始时即用此剂量,也可以安全地逐渐滴定到此剂量(通常通过强制的滴定方案)。通常,应设置多个剂量组,通过试验获得剂量-效应曲线,以证明剂量-效应关系。

一般情况下,剂量-效应关系临床试验要求各剂量组的效应形成较完整的量效曲线,量效曲线一般采用曲线拟合的方法获得,拟合的曲线应有统计学意义。一般不要求各剂量组间两两比较显示出统计学差异。对于中药复方制剂,由于剂量组设置相对较少,则应采用组间效应两两比较来确定量效关系。

一般而言,设置的剂量组越多、剂量梯度越合理,每组所需的样本量越小,反之所需的样本量则越大。此外也可以选择交叉量效研究、强制剂量滴定等试验设计方法进行剂量-效应研究。

(7) 析因设计 (factorial design)。析因设计是一种非常好且有效的研究设计类型,可以同时回答多个研究问题。从统计学计算上来看,这样的设计可以降低样本量。以中药与安慰剂对照试验为例,在各组基线水平可比的情况下,治疗方案分为辨证个体化治疗或标准化治疗方案,而同种治疗方案下中药剂型分为汤剂和胶囊。通过这样的设计就可以实现如下多种对比:①汤药与安慰剂汤药比较;②粉末胶囊与胶囊安慰剂比较;③个体化汤药与标准化汤药比较;④个体化胶囊与统一标准化胶囊比较;⑤统一标准化汤药与统一标准化胶囊比较;⑥个体化汤药与个体化胶囊比较。

这些数据对于验证中药不同剂型疗效差异是非常重要的。但是在比较中药与安慰剂对照的疗效差异方面,这种设计方法的把握

度就下降了,并且需要纳入更大样本量的受试者。

(8) 整群随机试验 (cluster randomized trial)。在整群随机试验中,相对离散的群体在参与试验之前进行整群随机,例如对社区全科医疗机构或临床病房进行随机分配。这种方法适用于不能或不适合对单个受试者进行随机分配,例如当对照组可能会因暴露于试验组干预措施而受到沾染,或者所需样本量巨大。在中药研究中,这种方法可以用于掩盖真药和安慰剂草药味的差异,避免破坏受试者盲法的实施。

(9) 技能型随机对照试验 (expertise-based randomized controlled trial)。技能型随机对照试验是将受试者随机分配到医师具有 A 技能的干预组或医师具有 B 技能的干预组,两组医师运用他们擅长的技能对受试者施行干预。这种设计可用于测试非药物疗效。

(10) Zelen 设计。Zelen 设计是 RCT 的变异类型,可以解决随机化所带来的问题(当患者随机分配后发现无法接受他们想要的治疗时,他们的依从性就会下降,同时有可能带来患者招募问题以及研究偏倚问题)。Zelen 设计的核心是在患者知情同意前就对患者进行随机分配,包括单组知情同意和双组知情同意。单组知情同意是被随机分配到治疗组的患者必须签署知情同意,知情后不同意的患者会被给予标准治疗。双组知情同意的患者首先被告知提供随机分配的治疗方法,如果患者拒绝,那么再调换成另一组(包括调换成治疗组)。只对保留原始分配的患者进行数据分析。Zelen 设计优势在于更容易招募受试者,并且有利于招募那些在治疗上有偏好的患者,患者纳入不受医生的倾向影响,样本更具有代表性,并且可以避免霍桑效应,主要应用于人群筛查以及患者对干预措施有明显偏好的情况。但同时 Zelen 设计也存在一定缺点,包括单组知情同意对照组未经知情同意而造成的伦理问题,因换组率增加而导致的结果稀释效应以及样本量增加等问题,在做统计时无法进行意向性分析(intentional analysis, ITT)。

在设计 RCT 时需要注意一些基本方法学要点,以下主要讨论知情同意、随机化方法、中药安慰剂的制备以及样本含量的估算。

(1) 知情同意。知情同意实施时,需要遵照受试者知情同意的标准操作规程,中医药临床试验在知情同意时也不例外,并且需要更加详细耐心的解释。在对患者知情同意的过程中,有 4 个要素不容忽视:必要信息、充分理解、完全自愿、书面签署。

①安排专业人员执行知情同意。知情同意最好由专人来进行,该专业人员大多以临床医生为主要组成,能作为接触到患者且最了解临床试验的人员,进行知情同意时,可以保证医患双方的利益,并可以解答患者提出的各种问题。

②知情同意的对象应具有自主判断力及行为能力。知情同意的对象大多为患者本人,但由于患者病情及行为能力等原因无法做出判断时,需采取一定措施,如无行为能力的受试对象,可由法定监护人代理;行为能力受限的受试对象,需法定监护人陪同在场;如患者和其合法代表(监护人)无识字能力,知情同意的过程应有见证人参加,经过详细解释知情同意书后,受试者或其法定代表人口头同意,见证人阅读知情同意书与口头知情过程一致,在知情同意书上签字,见证人的签字应与患者的签名在同一天。个别特殊的临床试验如不孕症的临床试验,有时候还需要患者的丈夫签字。

③口头交流内容应与知情同意书中的内容保持一致。试验开

始前,知情同意书须获得伦理委员会的审查批准。伦理委员会将对其内容进行逐一审核,并且在知情同意实施过程中不可以用其他版本代替。虽然不要求向患者口头告知时完全一字一句地照搬知情同意书的文字,但是对于关键问题的交代一定要严格遵照知情同意书的原文,比如盲法试验时对于安慰剂或盲法实施的解释等。口头介绍包括:研究目的、应遵循的研究步骤(包括所有侵入性操作)、研究持续时间以及可供受试者选择的其他方法等预期的受试者风险与受益,当受试者没有直接受益时,应告知受试者;参加研究是否获得报酬和承担费用情况;试验过程若造成损害时的保险、赔偿与治疗以及相关保密规定等。经知情对象初步了解试验信息及基本认同后,向受试者提供知情同意书,并逐一指出口头交流的信息在知情同意书中相应的文字,根据知情同意书补充口头交流内容。

④给予受试者足够的时间阅读知情同意书,并确保其能充分知情。研究者向受试者提供书面知情同意书副本,请受试者阅读知情同意书,还要给予他们足够的时间和机会阅读和询问研究的细节,以及询问其他任何问题,以便自主决定是否愿意参加研究。此外,还要确定受试者是否真正知情,可通过问答形式检测以确认其是否真正理解所提供的信息,以避免有些患者盲目服从或过度抗拒而不能正确判断。

⑤研究者获取受试者的知情同意签名。得到受试者同意后,应由受试者或其法定监护人在知情同意书上签字并注明日期,知情同意的实施者也需在知情同意书上签名并注明日期。一式两份签署后,向每一受试者提供一份双方都签过字的知情同意书。

(2)随机化及相关问题。随机化是 RCT 中去除混杂因素的最重要方法之一。随机的最初目的是为了确保持各组的基线因素均衡分配,以避免混杂因素更容易出现在某个研究组的情况。随机分配方案的隐藏应当做到使产生随机序列的受试者和进行随机分配的受试者分开,以最大限度地降低医生知晓分组后区别对待患者的可能性。在安慰剂对照试验中,可由独立的药房分发试验药物和安慰剂药包来完成随机。对开放性试验,最好使用电话或网络随机来实现随机隐藏。研究中心的随机化分配和药物配给集中由一个独立的机构或组织来安排和实施,整个试验的随机化方案由中央计算机统一控制,各个研究中心与这个独立机构通过电话、传真或计算机网络进行联系和临床试验的各种操作。基于电话传真的中央随机化系统(interactive voice response system, IVRS)和基于互联网的中央随机化系统(interactive web response system, IWRS)均依赖于计算机,使用时应充分考虑系统的可靠性和稳定性。设计完善的中央随机化系统一般应具有从受试者入组登记、筛选、随机、药物发放和紧急揭盲以及导出试验分组数据的功能。

①分层(stratification)。在小规模单中心临床试验中,应该重点关注疾病、干预措施疗效以及可能独立影响治疗效果的潜在的预测因素。因为对于大样本试验来说,混杂因素通常会在两组之间平均分配。而对于样本量较小的试验来说,特别是存在一些显著影响结果的变量时,应通过“分层”来确保这些变量在两组之间的平衡。此外,多中心临床试验的单个中心往往是独立的分层因素。由于参与同一项研究的受试者及其预后在不同的临床实践或医院条件下可能不同,因此确保每个研究中心分配到相似数量的受试者进入试验组或对照组是非常重要的。分层的另一个原因是平衡亚组中的受试者,以进行有用的亚组分析,例如不同的中医诊断模式。这将有利于测量哪个中医亚型比其他更多或更少地影响治疗效果,

它将成为分析试验以及在设计用于未来相关的研究的有用数据。该试验设计允许根据辨证进行个性化的中药治疗。在这种类型的试验中,受试者不仅被设盲而不知晓被施加了何种干预,也应该不知晓任何关于他们的辨证信息。

②整群(定群,clustering)。多中心临床试验的随机可能会出现各种各样的问题。同种治疗在每个中心的执行情况可能不同,或者同一个中心实施情况相同而不同中心之间不同,导致结局可能出现中心“聚集”效应。应尽一切努力使受试者招募、样本量和所有数据分析做到标准化,考虑可能存在的整群效应。因此可以考虑将每位参与这项研究的医生当作一个治疗“中心”,通过统计分析评估不同的医生对临床结果的影响程度。分析多中心临床试验时,对于理解和解释整群效应感兴趣的医生可以参考这种方法。

③协变量适应性随机化(covariate adaptive randomization)。不同的试验有不同的试验设计、不同的目的或需要考虑不同的预后因素,协变量适应性随机化也称为最小化法(或称中央随机化),它要求对协变量的当前状态进行不断更新,以达到影响临床治疗效果的预后因子的协变量平衡。适应性随机化无法在试验开始之前编制随机表,需根据已随机进入各组的受试者数量、影响疗效的预后因素的协变量均衡或之前分配给受试者的治疗反应成功或失败对当前受试者的分配概率进行调整,确定其分组概率。例如,在抗肿瘤药物的临床试验中,疾病的分期、病理分型、年龄等因素对治疗效果的影响很大,这时用分层随机化很难保证各组的例数和预后因素相接近且分析用病例数又足够。

(3)中医药安慰剂的制备。在临床研究中通常使用安慰剂实现盲法。安慰剂需要与被试验的干预措施在外观、颜色、气味、口味、包装、用法与用量方面保持完全一致,同时要经得起评价。在安慰剂的质量评价方面,目前还没有公认的评价方法。国外有学者提出了安慰剂质量检查清单(placebo quality checklist, PQC)的概念。该清单侧重于从盲法实施是否成功的角度对安慰剂的质量做出评价,但目前尚缺乏可靠的验证。国内有学者请从事制剂工作和临床研究的人员对安慰剂外观、颜色、口味等项目进行评分,以综合评分的结果作为考察的指标,为安慰剂的质量评价提供了有益的启示。

①中药安慰剂。中药由于其外观、口感、气味独特,制作与原试验药物完全一致的安慰剂,在临床工艺方面存在着较大的难度,尤其是汤药安慰剂的制作,由于保存期的限制,中药汤剂不可能一次性煎煮和包装,而是在原料配好后分批制作,即使在严格控制煎煮条件的前提下,仍存在较大的变异性,这对临床研究的操作实施提出了更高的要求,使用安慰剂对照时,要重视临床研究人员对安慰剂与干预措施的一致性评价。

为中药粉末制成的胶囊或药丸选择相应的安慰剂相对容易,可以使用不含有效成分的淀粉和着色剂来增强外观的仿真性。目前,可将汤药按照个体化剂量分装在密闭的塑料袋里,该技术促进了中药杂质安慰剂的发展,其味道与试验药物相似但无药理作用,使安慰剂外观更有说服力。这种方法的可信性已经通过试验得到证实。下一步工作是,用已经被确切证明了无效的食用色素和香料代替有杂质成分的中药安慰剂。

②针刺。虽然双盲安慰剂对照的 RCT 是评价干预措施特异性疗效的最优方案,但是安慰剂针刺的制作和实施一直都饱受争议。安慰剂针刺或称为“假针刺”(sham acupuncture),一般分为两类:刺

入性与非刺入性。刺入性安慰针刺是将针灸针刺入皮肤内(非常浅)或皮下(非腧穴或与疾病治疗不相关腧穴)而不进行手法操作。非刺入性安慰针刺是将针灸用一定的方法(如粘贴)固定于皮肤表面腧穴位置处并不入皮肤,或行针时退缩回针具的手柄内。但是,非刺入性安慰针刺几乎不可能对有针灸经历的受试者做到盲法。刺入性安慰针刺的设计存在明显的传统针灸理论的弱点。对于一项研究而言,安慰针刺刺入非穴或“不相关”腧穴一定不产生针刺效应是整个研究的前提假设,但是这个假设往往是缺乏依据而又难以证明的。

对于电针,安慰电针的设置相对明确。常用的方法有模拟电流法,即毫针刺入(或放置于)腧穴或规定部位后,针柄连接电极,电针仪指示灯亮,但没有电流通过针体。如果毫针刺法在电针组和安慰电针组都相同,仅仅是电流是否存在的差别,则比较的是通电与否的差异;如果两组在毫针刺法上也存在差异,那么两组所比较的就是整体层面的差异。

③太极拳。目前,太极拳 RCT 研究通常都未采用安慰剂设计,不能实现太极拳老师和受试者盲法。评价太极拳的特异性疗效,可以设计对照类型为不治疗(空白对照)。

④拔罐。由于操作方法和器械的特殊性,拔罐疗法目前还没有适宜的安慰剂对照措施。一方面,拔罐的作用机制(是单纯负压的作用还是结合热效应)还不是十分明确;另一方面,负压的作用是即时起效还是需要一段时间(至少 10 min)才起效,也不十分确定;而受试者的感觉比较鲜明,很难设置安慰剂对照而不破盲。虽然已经有研究尝试进行常规拔罐与小负压留罐对比的模拟安慰剂对照研究,但是尚未取得突破性进展。

因此,拔罐疗法作为一种技能相关的非药物疗法,与按摩、瑜伽等其他疗法相似,目前还是采用标准对照、空白对照、常规治疗对照等非安慰剂对照的形式进行研究。总之,对于受试对象和临床操作者来说,盲法较难实现。

(4)样本量估算。在 RCT 中通过样本量的计算纳入一定数量的研究对象,从而获得研究对象的重复观察,可以避免把个别情况误认为普遍情况,把偶然和巧合的现象当做必然的规律,通过一定样本量的重复,使研究结果具有稳定性,使假设检验达到预期目的。样本量的估计可以通过相应的公式计算获得,但是需要先确定基本统计学参数。

①确定单侧检验、双侧检验。当研究结果高于和低于效应指标的界限均有意义时,应该选择双侧检验,所需样本量大;当研究结果仅高于或低于效应指标的界限有意义时,应该选择单侧检验,所需样本量小。

②确定检验的第 I 类错误概率  $\alpha$ ,即检验水准或显著性。即假设检验第 I 类错误出现的概率,也就是假阳性错误出现的概率。 $\alpha$  越小,所需的样本量越大,反之越小。 $\alpha$  水平由研究者根据具体情况决定,通常  $\alpha$  取 0.05 或 0.01。

③确定检验的第 II 类错误概率  $\beta$  或检验效能。检验效能又称为把握度,为  $1 - \beta$ ,即假设检验第 II 类错误出现的概率,为假阴性错误出现的概率。检验效能即避免假阴性的能力, $\beta$  越小,检验效能越高,所需的样本量越大,反之就越小。 $\beta$  水平由研究者根据具体情况决定,通常取  $\beta$  为 0.2、0.1 或 0.05。

④处理组间差别  $\delta$  的估计,即确定容许误差。如果调查均数时,则先确定样本的均数和总体均数之间最大的误差。容许误差越

小,需要样本量越大。

⑤总体标准差  $\sigma$  或总体率。有效率越高,即试验组和对照组比较数值差异越大,样本量就越小,反之越大。

两样本均数比较,按下式计算。

$$N = 2 \left[ \frac{(Z_{\alpha} + Z_{\beta})\sigma}{\delta} \right]^2$$

$N$  为每组所需的例数。

两样本率比较,其样本量估计公式为:

$$N = \frac{[Z_{\alpha} \sqrt{2\bar{p}(1-\bar{p})}] + Z_{\beta} \sqrt{p_1(1-p_1) + p_2(1-p_2)}}{(p_1 - p_2)^2}$$

$$\bar{p} = (p_1 + p_2) / 2$$

公式中,  $Z_{\alpha} = 1.64$ ,  $Z_{\beta} = 0.84$ ,  $p_1$  与  $p_2$  分别代表试验组与对照组的事件发生率,  $\bar{p}$  表示两组发生率的平均值,  $N$  代表每组所需例数。

### 7.1.3 试验实施与资料收集

#### (1) 研究对象的确和招募

在 RCT 中,研究对象一般是患有某种疾病的患者,也有可能是无相关疾病的健康人。研究对象的选择要根据研究目的来确定,一般可以分为:①从该研究中可能获益最大且受害最小的人群,也是最易检出疗效的人群;②研究者特别关心的人群,如儿童、老人;③治疗效果不明确或可疑的人群。因此在选择研究对象时要明确病例的来源,包括来自哪一地区,哪一级医院,是门诊患者还是住院患者,研究疾病的诊断标准,病情的程度、疾病分期、有无并发症,患者的年龄、性别、病史、既往治疗史等。同时要根据研究要求制定出研究对象的人选标准和排除标准,在排除标准中应该列出不宜使用该药的情况,如心、肺、肝、肾功能不全者。因此,选择研究对象需要遵循以下原则:选择对干预措施有效的对象;选择干预对其无害的对象;选择能将试验坚持到底的对象;选择依从性好的人群;所有研究对象对研究知情认可,即了解研究目的、研究过程、可能的收益和危害等。研究对象必须签订《患者知情同意书》。

受试者招募过程主要包括招募受试者、筛选合格受试者、获得受试者的知情同意三个环节,以下是具体的受试者招募标准化操作流程。①确定受试者招募工作人员。可以是负责接诊的医生、分管的住院医师、护士,也可以是临床监管医生。招募人员确定后开始制定招募计划与需求。②确定招募方式。招募广告、专家电视健康咨询时进行招募宣传、社区义诊专家咨询会宣传以及邮件招募,最常用的形式是招募广告。③招募广告内容需包括:临床试验机构名称及地址、试验目的或试验概况、受试者主要纳入及排除条件、试验预期效益、受试者应配合事项、试验联系人及联系方式。招募广告不应包括的内容:宣称或暗示试验药品为安全、有效或可治愈疾病;宣称或暗示受试者将接受新治疗或新药品,而未提及该研究的试验性质;强调受试者将可获得免费医疗或费用补助;强调临床试验已经卫生主管机构或人体试验伦理委员会核准;使用名额有限、即将截止或立即联系等文字;使用含有强制、引诱或鼓励性质的图表、图片或符号等。④确定招募的场所。受试者招募的场所一般在医院和社区,以医院为招募场所可分为门诊、住院及健康体检中心,同时也可以社区卫生服务中心进行患者招募。⑤接待潜在受试者。设置专门的招募窗口由专人负责接待来访的受试者或者利用专线电话对咨询患者进行疑问解答。⑥按照纳入和排除标准以及医生的临床经验初步判断受试者入选可能性。⑦受试者合格性筛查。

对初步判断可能合格的受试者进行试验概况说明,进行体格检查或实验室检查,根据检查结果,再次确认受试者是否真正符合试验纳入排除标准。⑧受试者知情同意。对筛查合格的患者进行临床试验方案的详细说明,并告知其可能存在的获益和风险,做到充分知情。

(2) 干预措施的确立。RCT 干预措施的确立由研究目的决定。研究者首先必须针对干预措施的实施细节,做具体详尽的限定和描述,如药物的给药途径、剂量和用药时间,停药时间,严重不良反应出现时的处理原则,以及其他注意事项。同一治疗在不同用药条件下,效果和副作用的对比可能不同,研究中用药安排是未来实际用药的主要参考。

① 中药干预措施的确立。在实际的中医药干预研究中,中药具有非常重要的地位。一般而言,中药汤剂是中医辨证论治、个体化治疗中必不可少的干预方法,考虑到传统的中药煎煮方法非常耗时,并且具有强烈的味道,以及其对研究的依从性的影响。因此在实际的研究中,可以使用煮好的汤药按照每日剂量分装在密闭的、独立包装的袋子里,这样一方面保证了中药干预的标准化,同时患者不必自己煎煮中药,进一步增加了患者服药的依从性,并提高药物的一致性。同时当液体煎剂不适用时,研究者可以考虑从煎煮的汤药中提取粉末,制成中药片剂、颗粒剂或粉末胶囊等这些不同类型的中药制剂。

同时在试验过程中应当明确:a) 中药的具体成分,例如中药汤剂中各味药物的组成以及剂量,中成药的药名,生产批号,生产厂家;b) 中药的具体用法用量,例如药物的使用时间,使用次数,使用方法;c) 中药干预的实施周期,例如多少天为一个疗程,共治疗几个疗程;d) 同时治疗方案在 RCT 研究期间存在变化,也应当做出明确的规定;e) 中药干预的实施注意事项。

② 针刺干预措施的确立。针刺是一种基于施术者技能的操作性干预措施,除了严格执行经典的药物 RCT 的干预实施要求以外,针刺的 RCT 另具独特之处。

针灸师针灸技能差异。针灸师的受教育背景和临床工作时限的差别会导致针刺技能的差异。由于专业技能存在的差异,对于标准化处方(选穴、针法标准化)的针刺方案,必须在开始纳入患者前对所有参加试验的针灸师进行足够强度的针灸操作规范化培训,对于规定的穴位及针法进行详细的说明及演示,必要时利用针刺手法量化工具帮助实现针法的规范。培训的结果要求所有参加试验的针灸师在选穴及针法操作上能够实现一致性,即使个别针灸师对取穴和针法存在争议,也必须在给试验组患者施针时按照统一标准进行操作。对于允许加减选穴的 RCT,必须明确备选腧穴的取穴方法和针法。

针刺操作的精确安排。通常针刺 RCT 会对腧穴、留针时间等做出明确规定,如留针 30 min 或得气后即刻出针。有研究还明确规定行针时间,如每 5 min 行针 1 次。但是这些要求还不十分清晰。如留针时间起点从得气后开始计算还是从进针后开始计算?行针的时机是每刺入一个腧穴即行针还是所有的腧穴都刺入后再统一行针;刺入腧穴的顺序是否统一;从确保试验严谨性的角度,以上问题应有明确的规定。

③ 太极拳。由于太极拳干预措施的作用受到学拳方法、练拳方法、练习频次、周期、强度等因素的影响,而且太极拳老师的水平和学拳者的心态对太极拳干预措施的作用起着至关重要的作用,因

此,除了严格执行经典的药物 RCT 的设计和实施方式要求以外,需要考虑以下太极拳干预措施的特点。

太极拳老师的差异。RCT 试验中太极拳老师的背景差异,包括资历或从属的太极拳相关机构、习练太极拳的时间、太极拳教学经验,以及其他任何相关的经验,都将对太极拳教学效果产生影响。对于采用标准化太极拳教学方案的 RCT,如果涉及多位太极拳老师,必须在开始教学前使太极拳老师之间对教学方案达成一致,按照标准化的规范进行太极拳教学,包括讲解的语言、肢体动作演示的细节以及与太极拳学员的交流内容和方式等。这些都影响研究结果以及结果的推广应用。

太极拳方案的制定。试验设计阶段应当明确太极拳的学习方法,例如在太极拳老师指导下学习、自学、应用太极拳教学光盘辅助学习等,太极拳学习的步骤、强度,这会对太极拳的学习进度和太极拳干预效果产生影响;太极拳学习的内容,由于太极拳除了具体的肢体动作外,还有深厚的文化内涵、武术精髓以及哲学和医学相关理论,这些因素决定着太极拳练习者掌握太极拳的程度,进而也对太极拳干预的效果产生影响;太极拳练习的方法、单元次数、频率和疗程。如果练习单元的次数和频率在太极拳集体练习或太极拳课堂上练习以及自我练习或在家练习存在差异,应当做出明确的规定。如果干预的方案在 RCT 试验期间存在变化,也应当做出明确的规定。

④ 拔罐干预措施的确立。因为拔罐同属穴位刺激疗法,且同样具有非药物疗法的特性,拔罐疗法的研究在方法学上具有同针刺疗法相似之处。故而上述针刺疗法的某些方法学要点也同样适用于拔罐疗法。需要特别说明的问题还有以下几点。

拔罐操作者专业技能。与针灸师不同,罐疗师并非专门的“技术工种”,在考核医师资格时也没有专门考核拔罐操作水平的测试。故而,拔罐人员的水平参差不齐、较难量化。在进行拔罐疗法疗效评价时,如果不考核拔罐人员的技能水平,也难免会对评价结果产生不可预知的影响。拔罐治疗时罐内产生负压的大小、持续时间、作用部位选择等直接影响疗效的因素都与操作者的技术息息相关。因此,在研究设计、实施及报告的阶段,需要明确罐疗操作人员的资历。

拔罐刺激强度。与针刺研究不同,传统玻璃罐疗法在实施时无法以“刺入深度”来描述刺激强度,而拔罐疗法的刺激强度体现之一就是局部负压的大小。“负压”产生的大小依赖于罐疗师的熟练度和技能水平。解决这一问题的方法之一是应用真空压力表观测,随时注入或抽出空气来调节压力。然而,这一方法不适用于走罐、闪罐等要求手法操作的拔罐方法,所以,如何界定“负压大小”,如何使得研究中罐疗强度达成“标准化”,也是研究设计实施阶段的关键问题之一。衡量刺激强度的另外一个可以量化的指标是作用时间,针刺一般以 30 min 为标准作用时间,而不同拔罐疗法手法标准不同。临床上通常闪罐应用 20 次左右治疗,走罐应用 5 ~ 10 min/次治疗,而留罐时间根据留罐处表皮变化情况每次 5 ~ 30 min 不等。目前,鲜有研究针对留罐时间对疗效的影响进行比较和评价,故而,对于拔罐疗法的作用时间尚待确凿的临床试验数据来形成标准化的治疗方案。

(3) 结局指标的测量。对于使用定量方法测量结果的所有研究,中药疗效的相关结局应由独立的研究人员进行测量,如果可能,需要对结局评价者施盲,以尽量减少结局测量倚倚。结局测量也应在各组中采用同样的方式,以避免由于测量方式和测量时间不同对

测量结果产生影响。

临床试验结局的分类有多种,包括主要结局和次要结局、终点结局、替代终点结局、生活质量、主观结局和客观结局等。

主要结局指标的确定依据研究目的,国际上通常认为死亡、影响健康的重大医疗事件如中风以及受试者的生活质量是临床试验最主要的结局。而次要的结局指标通常包括症状或证候改善、疾病生物学指标(如血清学、病原学、影像学等指标),此外还有安全性结局和卫生经济学结局如成本-效果分析等。为研究设定主要结局指标非常重要,因为这可以避免在下结论时犯 I 型错误(即偶然出现的假阳性结果)的可能性。例如,一项临床试验使用 20 个结局指标,其中可能有一个指标在试验药物和安慰剂治疗之间在 5% 的水平上出现统计学的显著差异(即按 1:20 倍),因此这很可能被认为是一个显著性的结果。而事实上,如果其他 19 个结局指标的结果显示试验药物和安慰剂治疗之间没有差异,这也很可能仅仅是一个简单的随机事件。为避免出现这种随机事件需要预先确定主要结局指标。

在西医的临床试验中,大多数结局指标以疾病为核心。测量方式有多种:主观的、客观的和以受试者为中心的结局测量。根据中医药临床治疗的特点,临床试验采用主观结局指标可能更加适合,用来解释有临床意义的症状和体征的变化;这也包括中医证候的评价。而客观的结局指标,如实验室检查,目前少有研究并证实其测量与受试者相关的主观性结局变化之间存在相关性。

综合了不同健康状况维度的综合结局指标[例如健康评估问卷(health assessment questionnaire, HAQ)]具有将所有相关的结局整合成一个数值的优势。综合结局指标由于避免了因多重比较所致的  $P$  值调整而增加了临床试验的统计学效能,因此降低了对样本量需求。综合结局的劣势是其合并了结局和测量过程,并且合并时并未考虑其对应的临床意义。尽管对综合结局中各项指标的权重仍存有争议,但一般认为该结局指标能够反映患者的整体状态。但是,临床试验在报告综合结局时也必须同时报告构成综合结局的各项疗效指标单独的效应。

临床研究应当使用已经被验证的(理论上是国际公认的)测量结局。这样可使研究结果更易于推广,并且有利于试验被纳入系统综述和 Meta 分析。此外,今后还应当开发新的能够敏感捕捉到中医药角度重要临床变化的结局指标。例如,皮肤损伤后颜色的变化、月经周期规律的变化等,这些可能是反映中药疗效的重要指标。一些具体的结局指标在一定程度上得到证实后可与传统指标一起应用,以确保临床研究的结果与中医药的实践相关。

#### 7.1.4 统计分析

在 RCT 设计中,要根据试验目的、指标和方法、预期结果选择相应的统计分析方法。

首先在试验完成后,要按设计规定核对、整理资料,保证其准确和无遗漏。

其次,结果分析内容主要为两组或多组结局指标的比较、两组或多组剂量指标的比较、相关性分析及多因素分析等。最常用的一般原则和方法如下:计量资料数据,符合正态分布者用组间  $t$  检验,经方差齐性检验,方差相齐者用  $t$  检验,方差不齐者用校正  $t$  检验( $t'$  检验)。不符合正态分布者用非参数分析。当考虑治疗措施与结果的关系时,可以做相关性分析,如不同药量、不同疗程、不同剂量等因素与疗效的关系。也可以采用多元线性回归,获得研究因素对

结果的贡献。

计数及等级资料数据采用频数(构成比)进行统计描述,等级资料(如痊愈、有效、无效)组间比较采用秩和检验。计数资料采用  $\chi^2$  检验或 Fisher 确切概率法检验。如果考虑多种因素对结果的影响,可以采用多因素分析的方法,如 Logistic 回归分析,可以得到研究因素的净效应,并弄清有关因素的影响大小及方向。

#### 7.1.5 适应性试验

适应性设计是指在试验开始之后,在对试验的整体性与有效性不会造成破坏的前提下,依据前期试验所得的部分结果调整后续试验方案,从而及时发现并更正试验设计开始时的一些不合理的假设,从而减少研究成本,缩短研究周期的研究设计方法。适应性试验包括对样本量进行重新估计,反应变量-适应性随机化(response adaptive randomization),根据劣者淘汰(dropping-the-loser)原则选择干预措施,详细如下。

(1)通过样本量再估计。设计者在试验开始阶段重新设置参数,并根据试验前期获得的数据进行调整,以避免在试验终期,研究者获得全部数据时,发现因样本量不足而失败。因此,它避免了如试验初期对样本量估计不足等许多常见的缺陷,与下述的适应性试验联合应用,将更有利。

(2)反应变量-适应性随机化方法。其基于某个最优分配(optimal allocation)目标而进行,具有明确的最优分配目标,此目标往往依赖于一个未知参数的分布,故称为参数法。对于多阶段适应性设计,可以用前一个阶段的数据资料所得的统计量代替总体参数,计算下一阶段的待分配比例,再按照所得最优分配比例,采用一般的随机化方法对下一阶段的受试者进行分配;对于单阶段的适应性设计,可以根据已有数据所计算的最优分配比例获得下一位受试者分配到 A 组或者是 B 组的概率,从而在总的分配结果上达到分配的最优。通过反应变量-适应性随机化可对试验做出调整,以减少分配到治疗效果相对较差组的受试者数量。

(3)劣者淘汰原则或胜者优先(picking the winner)原则。指在试验实施阶段按照正式规定的决策规则更改随机分组,特别适用于在 II 期临床试验中选用(决定)不同剂量水平的药物。

### 7.2 队列研究

#### 7.2.1 定义

队列研究最早用于研究与疾病发生相关的病因或危险因素,将一群研究对象按是否暴露于某个研究因素分为暴露组和非暴露组,随访一定时间,比较两组之间所研究疾病或结局发生率的差异,以研究这项(些)暴露因素与疾病或结局之间的关系。

20 世纪 80 年代,队列研究开始用于医疗防治措施的评价,此时,暴露指具有预防保健或治疗作用的医疗措施,研究目的也从最初疾病发生、发展、死亡等转为治疗结局的评价。治疗性队列研究是指将特定患病人群根据其是否接受某种(类)治疗措施或接受不同类型的治疗措施分为不同的亚组,然后追踪观察一定时间,比较治疗组 and 对照组结局事件的发生率(如病死率)或治愈率的差异。

注册研究(registry study)和数据库研究(databases research)是近几年在中医药疗效研究中新生的队列研究,也称之为“真实世界研究”。数据库研究如台湾自 1996 年建立全民健康保险研究资料库(national health insurance research database, NHIRD),包含门诊、住院病历的中医、西医治疗方案、物理疗法、口腔科服务、处方药、医疗机构等登记信息,研究者利用这些数据库资

源设计回顾性队列研究来比较干预措施的疗效。

应用范围:在中医药临床研究中,可以通过比较不同的队列效果探讨中医治疗的优势病种、研究疾病的预后以及中医药的疗效比较效果研究(comparative effectiveness research, CER)等。

### 7.2.2 研究设计和方法学要点

按照研究结局在研究开始时是否已经发生,可以将队列研究分为前瞻性队列研究、回顾性队列研究和双向性队列研究 3 类。根据暴露因素(干预措施)与研究结局之间诱导期(从给予干预措施到发生预期研究结局之间的时间)的长短及预期研究结局发生率的大小,选择不同类型的队列研究方法。如针灸治疗某病的疗效评价,由于针灸起效快(即诱导期短)且效果明显(研究结局发生率高),所以适用用前瞻性队列研究;而观察药物的不良反应,一般情况下,由于上市药物不良反应的发生率较低,所以可采用回顾性队列研究。

(1)前瞻性队列研究是从现在时点开始,随访相当时期,比较两队列之间预期结局发生率的差异,以明确该暴露因素(干预措施)与预期结局之间的关系。随访过程可同期进行,并可选用最好的检测方法,按时记录,采用统一的判断标准,使两队列间具有很好的可比性。同时还可以观察其他次要因素对被观察对象预期结局的影响。在前瞻性队列研究中,根据被观察的人群不同,可以分为同群体队列研究和不同群体队列研究。同群体队列研究的被观察对象不论给予何种暴露因素(干预措施),都在同一个群体中,如同一个医院,同一个社区收集的观察对象。不同群体队列研究是指被观察人群的不同暴露因素(干预措施)不在一个群体里,但基本条件是相似的,如在一些中医院收集给予单纯中医药干预措施的研究对象较易,但收集单纯西医药干预措施的研究对象较困难,故也可选择同级别西医院,按统一的纳入标准,选择单纯西医药干预的研究对象作为对照组,比较两种干预措施的疗效。

(2)回顾性队列研究(历史性队列研究)的研究工作是从现在开始,而研究对象则是回顾性地追踪到过去某个时间作为始点,进入观测队列,也就是过去时点开始;研究对象的分组是依照当时的群体中是否暴露于某干预措施而制定的,研究的结局在目前开始研究时已经发生,从而探讨过去的暴露因素(干预措施)与目前发生预期结局之间的因果关系。该研究通常基于对既往医疗记录的数据进行分析,系当前“大数据”时代研究方法之一。该研究节省时间、人力和财力,观察药物不良反应或疗效的因果推断可采用此种研究设计。

(3)双向性队列研究是从过去时点开始直到现在,又从现在时点开始继续作同期随访到将来某个时期为止。这种方法在人力、财力和时间方面都可以节约很多,而且具有观察时间短的优点。

(4)病例队列研究(case cohort study)是将队列设计和病例对照研究设计相互交叉,融合两者的优点后形成的一种设计方法。该研究设计也称为巢式病例对照研究(nested case-control study)。病例队列研究是在队列内套用病例对照研究的一种设计,其研究对象是在队列研究的基础上确定的,以队列中所有的病例作为病例组,再根据病例发病时间,在研究队列的非病例中随机匹配一个或多个对照,组成对照组。但是其研究方法和分析方法仍与病例对照研究相同(主要是配比病例对照研究)。此种研究设计尤其适合于研究因素包括有复杂的化学或生化分析的前瞻性研究,也是数据库挖掘的方法之一。它只需收集那些被选为巢式病例对照研究的

对象而不是全队列的完整资料,从而减少了资料收集所花费的人力物力。其次,随着时间的推移,一项队列研究很可能要增加原设计中没有的某一暴露或混杂因素的内容,巢式病例对照研究能妥善解决这一问题。

(5)样本量计算。队列研究样本含量的估算有两个问题要考虑:一是对照组的样本量不能少于暴露组,两组相等时统计效率最高;二是需要考虑失访率,通常按 10% 估计失访率,即按计算出的样本量再增加 10% 作为样本量。样本的估算与 4 个因素有关:①非暴露组发病率( $p_0$ );②该因素引起的相对危险度的估计值(relative risk, RR);③希望达到的检验水准  $\alpha$  (通常取 0.05);④检验的把握度( $1 - \beta$ ,  $\beta$  通常取 0.20 或 0.10)。两组队列样本含量的计算公式如下:

$$N = \frac{[Z_{\alpha} \times \sqrt{2\bar{p}(1-\bar{p})}] + Z_{\beta} \sqrt{p_1(1-p_1) + p_0(1-p_0)}}{(p_1 - p_0)^2}$$

$$\bar{p} = (p_1 + p_0)/2$$

$$p_1 = p_0 RR$$

公式中,  $Z_{\alpha} = 1.64$ ,  $Z_{\beta} = 0.84$ ,  $p_1$  与  $p_0$  分别代表暴露组与对照组研究结局的发生率,  $\bar{p}$  表示两组发生率的平均值,  $N$  代表每组所需例数。

### 7.2.3 研究实施与资料收集

(1)确定研究问题。队列研究常用于因果关系假设和评价防治措施的效果。

(2)拟订研究计划。研究计划应当包括明确的研究因素及其特征,暴露(干预措施)的定义,组别的划分和研究对象的来源,样本大小及各组人数的比例,调查分析方法、内容和时间以及可能发生的偏倚及其控制方法,调查人员的培训,器材准备和预期结果等条目。

(3)选择研究人群及研究场所。由于队列研究要求的样本量较大,通常需要在几家研究场所进行观察,如不同的医院或社区,应制定相应的纳入和排除标准。对象的选择可以采用问卷或访谈、体格检查、中医辨证分型(针对中医研究)和实验室检查等方法获取基线信息。

(4)队列的形成(分组)。根据纳入和排除标准确定受试者是否暴露或是否接受某种治疗,暴露组(干预措施组)和非暴露组(对照组或其他治疗组)或暴露程度不同的亚组(如肿瘤受试者接受不同程度的中医药治疗)。

(5)测量暴露和随访(结局)。根据研究计划中确定的结局指标对观察对象进行定期测量、记录,采集的信息主要包括暴露情况(注意队列迁移的问题)和结局。对于治疗观察,应追踪和登记接受治疗措施的情况及治疗措施接受程度;登记所研究疾病结局的发生日期和测定日期,登记随访对象的迁移、外出及返回等信息。由于随访时间较长,研究小组应当制定提高受试者依从性的措施,包括建立研究受试者就诊的绿色通道,定期提醒系统,提供宣教手册,减免医疗费用,甚至为受试者提供随访的交通补助。

(6)队列研究不仅收集与暴露有关的资料,还收集与结局有关的资料。研究方案中必须明确定义暴露与否或暴露不同水平的测量,以及研究结局事件、观察终点及终止时间。暴露或结局事件的测量一定要准确,标准要统一。对于可能的混杂因素也要收集相关资料。

### 7.2.4 统计分析

固定队列和动态队列资料的统计分析方法不同。固定队列的每个个体都从同一时点开始观察,到研究规定的终点时间,在控制混杂因素后,对结局指标定期测量,比较两组或多组间该指标的差异,计量资料采用重复测量方差分析,计数资料采用生存分析;但对于动态队列,在任何时间点都会有个体退出或(和)加入,除常规的统计分析方法外,还可采用分层分析和多变量分析控制协变量,如 Logistic 回归和 Cox 比例风险模型,不仅可以控制混杂因素,而且可同时探讨多个暴露因素与研究结局的关联,及多个因素之间的交互作用。当队列研究样本量较小时,可能会出现某个暴露组人员过少的问题,也可能出现基线不均衡的情况。当这种情况出现时,研究者没有更好的办法来弥补,只能通过统计方法对基线不均衡的情况加以校正,甚至是进行配对。队列研究所采用的统计分析方法不同于一般的临床试验,除校正分析之外,亚组分析、多元回归分析、倾向指数评分等都是常用的统计分析方法。

队列研究中干预措施与疗效之间的关联强度一般用相对获益(relative benefit, RB)表示。干预措施有多个水平时,剂量-反应关系可以增加与疗效的关联强度。

### 7.2.5 注册研究

临床注册研究(registry study)是近年来国内外医药学界和政府部门都日益重视的一种新的临床研究形式。文献报道中也有学者译为“登记研究”。它不同于在同质人群中进行的临床试验,其具体定义为:有组织、有计划地使用观察性研究方法收集统一的临床及临床相关数据,对某一特定疾病状态下的人群,或患有某种特殊疾病的人群,或暴露在某一特定因素下的人群,进行特定结局的评价,从而达到预定的科学、临床和政策决策目的的一种研究形式。设计和实施良好的注册研究可以提供有关临床实践、受试者结局、安全性以及成本效益在现实世界的情况,促进证据的发展和运用,达到医疗保健科学决策的目的。注册研究属于真实世界研究(real world research, RWR)中一种重要的研究手段。

注册研究的开展多依托于电子信息系统,多数情况需要多中心及多方合作。其最大的优点在于可以在很短的时间内,将某个领域内的相关数据集合起来,有时甚至可以在全球医疗资源范围内对数据进行整合,这些海量的数据(大数据)可以为医学研究提供有价值的第一手临床资料。因此,注册研究具有其他类型的临床研究所不可比拟的优势。

注册研究属于观察性研究方法,可根据不同的研究目标和内容选择不同的设计方案,既可作前瞻性,也可作回顾性的研究。但是,注册研究最常见的研究手段还是前瞻性队列研究,不过与传统的队列研究相比,注册研究的手段更为灵活,比如一项注册研究可以同时为多个研究目的收集数据,且收集数据的方式有多种等。

注册研究的设计首先应该考虑登记的主要目的。此外,需要考虑的关键要素包括:(1)明确阐述研究问题,关注的临床问题将直接指导登记的设计,包括暴露和结局的选择,和目标人群的界定;(2)根据研究目的和研究问题选择适合的研究设计类型(如前瞻性队列研究、病例对照研究、结局研究等);(3)明确目标人群、诊断、纳入和排除标准,登记人群应尽可能地接近目标人群的特征,因此尽量保证受试者的代表性(即减少排除);(4)确定是否需要对照组,以描述性研究为目的的注册研究可不设置对照组,以分析性研究为目的的注册研究可依目的不同选择内对照、外对照和历史性对照;(5)选择临床上有意义的、并且与受试者及医疗决策相关的

临床结局;(6)确定数据来源,选择最有效和最可靠的数据收集方法;(7)确定随访时间及样本量,纳入的研究对象数量和随访的时间和登记的整体目标一致。

需要注意的是,注册研究毕竟属于观察性研究方法,因此会存在观察性研究所固有的问题,比如已知或未知的混杂因素比较多,组间可比性差等。因此,注册研究的结果必须要排除各种各样混杂偏倚的影响,这除了在研究设计阶段对登记设计进行审查,评估偏倚的可能来源之外,方法学专家也在方法学上进行了多方尝试,如工具变量、倾向性评分、多因素分析模型等。因此,这类研究通常需要引入方法学和统计学专家参与设计和数据分析。

## 7.3 病例对照研究

### 7.3.1 定义

病例对照研究属于临床流行病学的观察性研究方法,属于因果关联推论的一种分析性研究,是将现在确诊的患有某种特定疾病的患者作为病例,以未患该病但具有可比性的个体作为对照,通过询问、实验室检查或复查病史,搜集既往各种可能的危险因素暴露史,测量并比较病例组与对照组中各因素的暴露比例,经统计学检验,若两组差异有意义,则可认为因素与疾病之间存在着统计学关联。经典的病例对照研究主要用于病因推论。

目前,也有学者将该方法从病因与危险因素研究逐步扩大到疗效评价。此时,研究对象的临床结局(如治愈和未治愈,好转和无好转)成为分组的依据(而不是患病情况),既往的暴露因素为接受的治疗措施(而不是既往暴露的危险因素),通过比较两组不同结局患者的既往治疗措施的不同,推论既往的治疗(暴露)和结局(病例)之间是否相关。

### 7.3.2 研究设计和方法学要点

病例对照研究的设计原理是:首先根据一定的条件将某一个人群中患有研究疾病的人群作为研究对象,组成病例组,并为每个病例选取一定数量的研究对象作为对照组;对照组均产生于该人群内部,属于其对应的病例发病时尚未发生相同疾病的人,并且按年龄、性别、社会阶层等因素进行匹配(即危险集抽样),然后分别抽出病例组和对照组的相关资料及生物标本进行检查、整理,最后按病例对照研究(主要是匹配病例对照研究)的分析方法进行资料的统计分析和推论。该研究的研究步骤及设计要点如下。

(1)明确研究目的。研究目的是确定设计步骤的出发点,确定本次研究所要分析的结局(终点事件)、暴露(干预措施)、协变量等。

(2)确定结局组。当病例对照研究用于疗效评价时,分组依据不再为是否患病,而是是否出现某个临床结局,因此,在本节中采用“结局组”和“对照组”的说法,而不是原来的“病例组”和“对照组”,目的是为了避免读者产生混淆。结局指的是研究者预期的结果事件,可以是定性的指标,如治愈和无效,也可以是定量的指标,如某个血清指标的水平。对于结局的确定需要采用国际或国内通用的标准,并对结局组和对照组使用统一的判断标准。

(3)确定对照组。即在患该病的人群中选择一定数量的,尚未发生临床结局的病例,按年龄和性别等因素与结局组进行匹配。对照的来源可以有人群对照和医院对照两种模式。人群对照的代表性比较好,但人力、物力花费较大,且应答情况较差。医院患者对照合作性好,但是容易出现选择性偏倚。事实上,并不存在完全理想的对照,最好的办法就是设置多组不同类型的对照,而这又会增加

研究成本和分析的复杂程度。现实中就是不断使各种矛盾处于最佳平衡,这是病例对照研究的一个难点,也是容易产生偏倚的环节。

(4) 样本量计算。目前,尚未查到用于疗效评价的病例对照研究样本量特殊的计算公式,可以参照经典病因学推断时病例对照研究的计算公式。病例对照研究的样本量与 4 个因素有关:①对照组中的估计暴露率( $p_0$ );②估计具有意义的最小增加或减少的比值比(OR);③希望达到的检验水准 $\alpha$ (通常取 0.05);④检验的把握度( $1-\beta$ ,  $\beta$  通常取 0.20 或 0.10)。可以通过以下公式来估计。

$$N = \frac{2\bar{p}(1-\bar{p})(Z_{\alpha} + Z_{\beta})}{(p_1 - p_0)^2}$$

$$\bar{p} = (p_1 + p_0)/2$$

$$p_1 = p_0 OR / [1 + p_0 (OR - 1)]$$

公式中,  $Z_{\alpha} = 1.64$ ,  $Z_{\beta} = 0.84$ ,  $p_1$  与  $p_0$  分别代表病例组与对照组的暴露率,  $\bar{p}$  表示两组暴露率的平均值,  $N$  代表每组所需例数。

### 7.3.3 研究实施与资料收集

既往暴露情况的收集:在疗效评价的病例对照研究中,暴露情况指的是既往的治疗或干预措施。每一项干预措施(如药物、手术、患者教育等)都需要一种独特贴切的暴露确定方法。当干预措施是一次性行为时,如手术,其测量只需要确定是否暴露及何时暴露。当干预措施要持续一段时间时,如药物,其测量需要综合考虑剂量、干预频率和持续时间,以确定暴露的强度。有时,还需要考虑干预所处的环境。此外,暴露的测量还需要综合考虑暴露和结局之间的理论/生物学关联,暴露导致结局事件的诱导期和潜伏期,及暴露状态是否随时间变化等,最终确定暴露的测定方法和测量时间。总之,暴露的测量不存在“万能”的解决方案,需要根据具体问题具体情况来确定。

对于中医药的疗效评价,所涉及的暴露因素既包括具体的、单一的暴露因素,如中药注射剂、太极拳、穴位贴敷等,也包括抽象的、复杂的暴露因素,如辨证论治、中西医结合治疗或多种方法联合使用。相对于具体的暴露因素,抽象的暴露较难进行定义和测量,没有国际国内公认的定义和测量方法,目前可以采用的方法有文献研究法、头脑风暴法和德尔菲法。有的学者选择将中医药整体作为一个暴露因素,这时要精确定义暴露程度——即中医药的不同暴露水平,例如可分为以下几个类别:只接受规范中医药治疗(极高暴露);西医治疗基础上,在中医医院接受长期连续中医辨证论治综合治疗,包括中药、针灸、气功等治疗(高暴露);西医治疗基础上接受长期中成药治疗(较高暴露);西医治疗基础上给予短期或间断的中成药治疗(弱暴露);纯西医治疗(无暴露)。或者可以把中医药各组成要素作为独立暴露因素,要在调查表中详细罗列。这种做法在事先需要做大量工作把可能的暴露因素全部鉴定出来。

### 7.3.4 统计分析

病例对照研究中,暴露与结局之间的联系可以用 OR 值和 OR 值的 95% 可信区间进行估计,此外,作为观察性研究,病例对照研究无法像 RCT 那样,通过随机化一劳永逸地平衡所有暴露因素(或者研究要评价的干预措施)以外的混杂因素(例如,人口学资料 and 生活方式等)。因此,有必要采用多因素 Logistic 回归方法来校正混杂,结合临床知识,通过将可能的混杂因素纳入模型,可以得到暴露因素调整后的 OR 值。其中,对于匹配的病例对照研究,采用条件 Logistic 回归分析暴露与结局的关联。

### 7.3.5 巢式病例对照研究

在比较效果研究中,病例和对照常嵌入队列研究之中进行,即形成巢式病例对照研究。巢式病例对照研究,又称嵌套式病例对照研究或队列内病例对照研究(case-control study nested in a cohort),是将病例对照研究和队列研究进行组合后形成的一种新的研究方法,即在对一个事先确定好的队列进行随访观察的基础上,再应用病例对照研究(主要是匹配病例对照研究)的设计思路进行研究分析。巢式病例对照研究的优点较多,比如:相比于队列研究,可以节约人力、物力和财力。此外,由于该研究中的暴露资料(即治疗措施)是在结局发生前收集的,如果研究结果显示治疗措施与临床结局存在关联,那么该关联与因果推断的时间顺序相符合,因果联系的推断更有力。最后,巢式病例对照研究的统计效率和检验效率也高于病例对照研究,还可以计算临床结局发生率。

巢式病例对照研究的设计原理是:首先根据一定的条件确定某一个人群(一般为患某种疾病的人群)作为研究的队列,收集队列中每个成员的有关资料信息和(或)生物标本,对该队列随访一段时间,将发生在该队列内的出现临床结局(如治愈、好转等)的病例全部挑选出来,组成“结局组”。然后,为每个“结局”选取一定数量的研究对象作为对照组。对照组均产生于该队列内部,属于其对应的尚未发生临床结局的患者,并且按年龄、性别、社会阶层等因素进行匹配(即,危险集抽样),然后分别抽取事件组和对照组的相关资料及生物标本进行检查、整理,最后按病例对照研究(主要是匹配病例对照研究)的分析方法进行资料的统计分析和推论。

### 7.4 病例系列研究

#### 7.4.1 定义

病例系列研究是对单个病例报告的集中描述与分析,一般包含 10 个以上病例的详尽临床报告,包括临床表现(症状、体征和实验室检查结果)、治疗、治疗后的反应及结局,是作者对多年积累的病例的总结。其目的在于通过探讨一组研究群体的详细临床资料或病史记录,进行观察、分析干预措施与结果之间的关联。

#### 7.4.2 研究设计和方法学要点

病例系列研究基本的设计模式是,在患所研究疾病的人群中挑选出符合纳入标准的接受某种干预措施的受试人群,在一定观察期后观察某种既定结局指标的情况,探讨干预措施与结局之间可能的关联。其设计要点包括:仅对接受相同干预的一组受试者的临床表现、临床治疗及疗效情况进行描述和评价;未设立对照组;分析结局与干预措施的关系,提供的证据只能为经验型证据,不能作为因果联系的最终结论。

从新发现的异常病例中提出某种可疑假设,异常病例包括出现的频率和分布异常、临床表现异常、实验室检查结果异常及对治疗的反应异常等,研究者可分析出现异常的原因。当然,在此阶段,只是根据已有的理论或很少的相关信息提出假设。因此,病例系列研究可以用于观察特殊疾病(AIDS、非典型肺炎等)、罕见慢性病、暴露与结局时间较长的研究、并发症和不良反应;可以用于因伦理等问题无法实现的对照研究设计;也可以用于观察临床对照试验排除的患病人群,为将来进一步试验研究提供依据。其优点还在于费用低廉、容易进行、可以提出假说,为未来指明研究方向。然而,病例系列研究由于不能提供因果关系解释,因此证据强度较低;其结果基于所选择的受试者,因此将该结果用于其他受试者的可应用性不明确;另外,病例系列研究在设计时通常未考虑很多潜在的混淆因

素,故容易高估观察疗效,因此在采用该研究报告的结论时要非常谨慎。

病例系列研究适用于如下范围:(1)疾病预后清楚、受试者有明显的选择倾向、无其他可用的或可接受的治疗方案;(2)罕见病、特殊病或者研究周期较长的疾病危险因素、预后、疾病演变(自然史)等问题的研究;(3)被其他研究设计排除在外的特殊人群;(4)特殊新药、新疗法疗效及药物不良反应监测等。

按照时间顺序,病例系列研究可以分为回顾性病例系列研究和前瞻性病例系列研究。回顾性病例系列研究最常见,撰写于治疗结束后,治疗实施前未进行文献回顾,也无既定的治疗计划。虽然能反映出经治医生最佳医疗水平,但可能由于实施治疗时未能采用最佳结局指标而导致病例报告可信度降低。前瞻性病例系列研究需要临床医生事先对感兴趣的医疗问题进行文献研究,继而确定医疗方案和结局评价指标,等待符合要求的受试者前来就诊。当受试者来诊时,按照事先既定方案进行治疗,并在治疗前、中、后期测量结局评价指标。此种报告在实际看到受试者前就制定医疗方案,虽然这一点面临着一定的质疑,但是其优点在于需要相对较短的撰写报告的时间,也可能会因为事先准备医疗方案而提高医生的医疗水平。

按照结果对照的类型,病例系列研究还可以分为仅有治疗后结果的病例系列研究和无对照、治疗前后比较的病例系列研究。前者指结局指标仅仅在干预措施之后被记录下来,并且无同期平行对照组可以比较,是为了探讨影响受试者结局的因素,通过对受试者已经出现的结局,调查是否因接触某致病因素或接受某治疗措施而提出一系列叙述性资料。该设计不设对照组,有时可采用本人或他人既往研究结果作为对照,也称为历史性对照。而无对照治疗前后比较的病例系列研究指研究对象的结局指标在接受干预前和后各测量一次,便于比较。这种方法在临床干预性研究中较常应用,此外也用于某些新药Ⅱ、Ⅲ期临床研究中评价药物的安全性和临床疗效。

病例系列研究常见的偏倚有:由于对过去事件或经历回忆的准确性和完整性不同造成的回忆偏倚;由于非随机抽样,造成选定研究对象与未被抽取的人群在某些特征上存在系统差异而出现的选择性偏倚;由于杂志编辑或作者过分重视“阳性结果”,而造成有统计学差异结果的文章比没有统计学差异结果的文章被发表可能性更大的发表偏倚。

#### 7.4.3 研究实施和资料收集

研究实施的过程是首先围绕研究目的,制定完善的研究方案、临床研究标准操作规程及病例报告表;然后根据具体研究内容,进行合适的样本量估算;制定严格的诊断、纳入与排除标准,确定合适的受试者;在观察期内根据制定的病例报告表全面、完整地收集数据;研究人员独立、客观、清晰地对结局进行评估;最后评价研究结果的临床价值为进一步研究提供有价值的线索。需要强调的是,病例系列研究属于观察性研究,研究者不能强行对所观察对象实施干预。

对于研究对象,设计研究时要充分考虑对病种的选择。例如,自愈性疾病必须有对照,但如果是比较罕见的或者慢性疾病、符合“全”或“无”规律的疾病[“全”是指在采用此种治疗方法之前,全部受试者都会发生某不良结局(如死亡),而采取此种治疗方法之后,部分受试者生存下来;“无”是指在使用此种治疗方法之

前,部分受试者因病死亡,而使用此种治疗方法之后,受试者未因该病而死亡],或者涉及伦理学问题而不能设置对照时可以使用病例系列研究的方法。由于研究对象的选择并未采取随机的方法,因此不能仅选择疗效好的病例,避免人为夸大疗效。如果是前瞻性设计,应该连续招募病例;如果是回顾性设计,应该选择某时间内接受该种疗法的全部病例。对于受试对象的诊断、纳入和排除标准的制定与其他研究相似,然而,病例系列研究的样本量估算并没有统一的标准或公式,目前中医药的病例系列研究并没有可供参考的样本量计算公式,建议参考横断面研究的样本量估算方法。

对于干预措施,病例系列研究虽然不要求将每一位受试者的信息都详尽描述,但是中医药病例系列研究应该体现中医药的特点,如描述辨证分型依据和加减原则、中药的组成、来源、制剂的质量、剂型、给药途径、剂量和疗程,针刺及拔罐的种类、手法操作细节、穴位选取的方法等,以便其他中医师能够结合临床实际借鉴或使用。

对于结局指标,病例系列研究结局指标的选择主要依据研究目的制定。国内外中医药病例系列研究结局指标多选取疗效指标和安全性指标。治疗性的中医药病例系列研究应选择医生和受试者最为关心的临床事件作为结局指标,尽量选取可以获得的终点结局指标,在无法获得终点结局指标时可以选择替代结局指标(中间指标),包括受试者的生活质量。另外,也要注意观察受试者对治疗的依从性等指标,借以判断干预措施实际应用的可行性。

#### 7.4.4 统计分析(表2)

根据病例系列研究的设计模式,也可以列出如下四格表,将结局填入相应的表格内,对暴露或接受某种因素的结局进行分析和比较。

表2 病例系列研究的结果分析

		结局		合计
		+	-	
暴露或接受某干预因素	是	a	b	a + b
	否	c	d	c + d

与有对照的研究不同,对照组的资料c和d可以是来自历史性对照的资料,也可以是该组病例治疗前的数据资料。二者不同的是,对于结局指标为计数资料的数据来说,历史性对照的统计分析可采用一般的四格表卡方检验(或根据数据特征选择替代的Fisher精确概率法、秩和检验等),而自身治疗前后数据对照的统计分析可采用配对四格表的卡方检验。对于计量资料类型的结局数据来说,历史性对照的统计分析可根据数据分布类型选择两个独立样本t检验或独立样本秩和检验的方法进行统计分析,而自身治疗前后数据对照的统计分析可采用配对t检验或配对秩和检验的方法进行统计分析。当考虑分析观察结局与时间序列的关系时,可采用生存分析等方法进行统计分析获得干预措施的生存率或者某些事件的发生率与既往研究结果的比较。

#### 7.5 卫生经济学评价

卫生经济学研究是应用一定的经济分析与评价方法,将相关卫生规划或卫生活动的投入和产出相联系进行比较评价的研究。对于药物和医疗卫生手段的研究,在疗效和安全性确定的情况下,开展卫生经济学评价对于卫生和医疗决策的制定是非常关键的,卫生经济学评价结果可减少或避免可能造成的损失或浪费,使有限的卫生资源得到合理的配置与有效的利用。在过去的几十年中,关于补充替代医学的研究大量开展,如效力、作用机制以及关于药材和药

物产品质量标准化的研究等。研究者们主要关注高负担疾病,如慢性病的预防和管理,在这些疾病中,西方医学进展迟缓且无确切疗效证据,而补充替代医学和西方医学的联合应用则逐步得到推广。因此,目前越来越多的中医药临床研究在进行疗效评价的同时,进行经济效益评估;即除了在研究过程中,收集疗效和安全性数据,还在试验中获取成本数据。卫生经济评价讨论与分析的中心是选择,它的基本任务就是确认、衡量、比较与评价候选方案的成本和获益。

### 7.5.1 卫生经济学评价方法

卫生经济学评价方法主要有四种:最小成本分析(cost-minimization analysis, CMA)、成本-效果分析(cost-effectiveness analysis, CEA)、成本-效用分析(cost-utility analysis, CUA)和成本-效益分析(cost-benefit analysis, CBA)。

(1)CMA。CMA 是在效果(效用,效益)相同的基础上,对不同药物或医疗手段的费用(成本)进行比较,即对效果(效用,效益)相同或者相近的两个或多个备选药物的成本进行比较,从而选择成本最小的药物的分析方法。如普通药品和名牌药品的比较,若普通药品和名牌药品的疗效相同,普通药品的成本一般会低于名牌药品。

采用 CMA 需要备选方案的疗效相同,因此在药物经济学评价研究时需要证明两者的疗效相同。但是,由于统计学上采用反证法,统计结果不支持疗效有差异并不等于支持疗效相同,因此很难判断两种或多种药物是否疗效相同。再者是统计学支持疗效相等并不等于临床疗效相同。正是由于疗效相同判定的困难,CMA 有时会产生误导作用。由于 CMA 严格的限制条件,在药物经济学中应用范围较局限,但是 CMA 只需对比各备选药物的成本大小即可,计算简单,评价结果也易于理解。因此在疗效相同或者相近的情况下,CMA 是进行药物经济学评价的首选方法;而如果疗效不同,则应该采用其他分析方法。

在应用 CMA 分析方法时,如果有药品不良反应发生,成本核算中应纳入处理不良反应的成本。

(2)CEA。CEA 是较为成熟的药物经济学评价分析方法之一,是将药物治疗的成本与所产生的效果进行分析,进而对两个或两个以上备选药物的成本和效果进行评价,根据评价结果评估药物的经济性。效果即为使用某种或某些药物后的个人健康产出,这些产出表现为健康的结果,如治愈率和发病率的下降等。成本效果分析要求比较的效果单位一致,但不需要疗效相同,且结果易于为临床医务人员和公众接受,是药物经济学研究中应用最广泛的药物经济学评价方法之一。

成本效果分析主要采用两种表示方法。

①成本效果比。即平均每产生一个单位效果所需的成本。其计算公式是:成本-效果比(C/E)=成本/效果。

②增量成本效果比(incremental cost effectiveness ratio, ICER)。其是指相对某种药物,检验药物每增加 1 个单位效果所增加的费用。其计算公式是:ICER=(A 药治疗成本  $C_A$  - B 药治疗成本  $C_B$ )/(A 药效果  $E_A$  - B 药效果  $E_B$ )=A 药与 B 药的成本之差  $\Delta C/A$  药与 B 药效果之差  $\Delta E$ 。

CEA 的产出采用非货币形式的健康产出,但是能够反映效果的指标有很多种,因此选择合适的效果指标至关重要。我国的临床效果评估研究,由于国家医疗监测系统不够完善以及受试者的流动性大等原因,很难获得长期临床效果指标,大部分情况下使用的是有效率、治愈率等指标,这类指标既不能反映受试者的长期生存状况,而且评判时容易受到结果观察者的主观影响,尤其是在没有执

行盲法时,结果偏差较大。

应用 CEA 进行分析时,采用终点结局指标还是中间结局指标,不良反应的体现差别较大,尤其是长期的不良反应。那么可能只有终点结局指标才能反映出来,而中间结局指标则很难体现。因此,采用中间结局指标可能无法体现不良反应的影响,而只有采用生命年等长期效果指标才能够反应出不良反应所造成的影响。

(3)CUA。CUA 是比较成本和经质量调整的健康效益产出量,来衡量药物或治疗措施的一种经济学评价方法。CUA 是 CEA 的发展,与 CEA 有许多相似之处——两者均用货币来衡量成本,并且测量结果也都采用临床指标作为最终结局的衡量参数。所不同的是产出的测算,CEA 的产出效果采用与治疗目标一致的自然单位,如延长寿命时间、增加体重、降低的血压值等。而 CUA 的效用评价采用健康改善指标,效用指标不仅关注生命的数量,同时还考虑了生存活量的因素。CEA 关注生命是否延长,而 CUA 更关注延长的生命年质量,并采用生存质量调整年(Quality Adjusted Life Years, QALY)等指标作为效用参数进行分析。由于 CUA 测量生命的质量和数量,因此能够更好地反映慢性疾病,尤其是肿瘤等器质性病变的治疗效果。同时,由于 CUA 采用同一结局指标,能够方便地比较 CEA 所不能比较的结局指标不同的药物,并且能和以前所做的 CUA 进行比较。CUA 的计算方法和 CEA 的相似,有成本效用比和增量成本效用比两种方法。

CUA 在药物经济学评价中有着广泛的应用,也是药品不良反应的良好评价工具。CUA 在中医或中药评价中具有明显的优势。中医药的治疗目标之一是改善症状,而症状的改善则可以提高慢性疾病受试者的生活质量,但是很多治疗效果指标如血压、血糖等可能不能够反映受试者的临床结局。关于中药配合介入治疗肝癌的临床研究表明,中药配合介入治疗组与单纯的介入治疗组比较,瘤灶疗效无明显差异,但是中药配合介入治疗组的受试者的生活质量(卡氏评分)和生存时间(随访时间为 4 年)均优于单纯的介入治疗组。因此,中药与常规治疗慢性病的 CEA 结果可能没有差别,但是由于中药改善了受试者的生活质量,根据 CUA 结果,中药治疗优于常规治疗。

(4)CBA。与 CEA 不同,CBA 分析中,结果以货币形式表现,因此 CBA 是一种将成本和结果均以货币为单位进行测量的经济学评价分析方法。CBA 可以比较不同药物对同一疾病的治疗效益,还可以比较同一药物对不同疾病的治疗效益,甚至比较不同药物对不同疾病的治疗效益。它是一种具有政策导向性的经济学评价分析方法,其结果对于制定卫生政策如药物报销指导目录等具有指导意义。然而,许多临床效果指标,如患病率、病死率和生命年等很难以货币单位来进行衡量,同时临床医疗人员和公众难以接受以货币单位来衡量生命和健康的价值,因此,CBA 在药物经济学研究上的应用相对较少。

### 7.5.2 卫生经济学研究中的成本测算

药物经济学研究 4 种分析方法的不同点是结局测量各异,相同的是成本测量一致,即无论采用哪种方法,均需要对成本进行测算。成本虽然可以根据客观指标来测量,但不同的研究角度则决定了成本测量的范围不同,同样决定了研究结果的外推性不同。以全社会角度进行研究则需要将直接非医疗费用如因治疗疾病产生的交通费和间接费用如因病误工损失的收入等纳入成本测算,而以医疗卫生提供者角度开展研究则只需要纳入直接医疗相关成本,可以不纳入直接非医疗费用。药物经济学评价通常的研究角度包括:全社会角度、“医保”付费方角度、受试者角度和医疗服务提供者角度等。

全社会角度对成本的测算是最为全面和详细的,其他角度的测算成本将包括全社会角度测算成本中的一部分或几部分。药物经济学研究中全社会角度成本测算主要包括以下几个部分。

(1) 卫生服务系统内的直接成本。直接医疗费用,如药物的治疗费用,使用药物产生的附加费用如注射器费用、住院费和护理费等。这部分费用可在医疗机构通过查询就诊记录中收集的费用信息或者从医院财务记录中获得。相对而言,直接医疗费用最容易收集,所获得的信息的可靠性也较高。但如果是受试者自行购药且自行应用,自行购药的这部分费用在医疗机构中没有记录,这时需要通过问卷调查或向受试者索要购药发票等途径获得。

(2) 卫生服务系统外的直接成本。直接非医疗费用,受试者为了寻求医疗卫生服务所产生的费用或消耗,如受试者的交通费、陪护费等。这些成本测算没有记录可查询,多需通过问卷调查获得。

(3) 间接成本。由于疾病、伤残、死亡给社会带来的经济损失,其通过劳动力有效工作时间的减少和工作能力的下降而产生。常见的间接成本测算方法有以下 4 种:①现值法,用工资标准来计算间接经济负担,即工资标准×因病损失的有效工作时间。②人力资本法,可以用人均国民生产总值或者人均国民收入来计算人年工作损失带来的间接经济负担。③意愿支付法,指为避免特定疾病个体所愿意付出的货币值,它认为人的价值包括人的健康和收入,通过效用函数估算,数据通常采用问卷调查获得。④磨合成本法,在完全劳动力市场条件下,由于受试者缺勤、休工需要聘请新的职工来代替,通过招募、培训后才能上岗,因此在这一段磨合期中除生产损失外,还需要有培训上岗成本的投入,以此来估计患病后对总生产成本造成损失的价值,这种间接成本比用人力资本方法估算要小。

(4) 隐性成本。隐性成本指受试者因病遭受的痛苦、悲伤、抑郁等难以用货币确切计量的成本。

由于治疗的时间跨度,以及对过去进行回顾性研究时,需要使用贴现率,即将不同时点的成本或者效益换算成同一时点的成本和效益。其计算方法如  $C = \sum_{t=1}^n \frac{Ct}{(1+r)^t}$ , Ct 为时间段内的总成本; r 为贴现率; n 为时间段的数目。并且,在对不确定因素进行分析时,往往通过变动贴现率指标进行敏感性分析。

### 7.5.3 卫生经济学评价注意事项

(1) 卫生经济学评价中不可忽视不良反应。据世界卫生组织(WHO)统计,各国住院受试者药品不良反应发生率为 10%~20%,其中 0.2%~2.9% 的受试者因为严重不良反应而死亡。在我国,有研究估计每年用于治疗不良反应的医疗费用高达 15 亿元。不良反应的发生不仅增加了受试者的治疗费用,体现在成本测算中(一般体现为成本增加),还对结局指标产生影响,体现在部分产出指标中(一般体现为降低产出)。药品不良反应造成增加的成本包括直接医疗费用、直接非医疗费用、间接费用和隐性成本 4 个部分或者其中的某几个部分。药品不良反应对产出的影响如不良反应带来的病死率/死亡率/致残率的增加,治愈率的降低, QALY 的降低和收益的降低等。在药物经济学评价中,不良反应对成本和产出指标的影响意义重大,对药物经济学评价研究结果正确合理指导临床和卫生政策的制定具有重要意义。但是,在药物经济学评价研究中,不良反应的报告情况并不完善。而忽略不良反应的药物经济学评价研究可能导致结果更有利于试验药。因此,药物的不良反应需要治疗时,必须药品成本和负产出,否则它可能改变药物经济学评价的结果。

(2) 卫生经济学评价中的不确定性因素。敏感性分析(sen-

sitivity analysis)是药物经济学评价中广泛采用的处理不确定性因素的一种方法。敏感性分析是评价经济学模型可靠性的一个过程。它通过几个主要变量在一定范围内的变动,分析检查对结果带来的影响。国内外各种指南均推荐采用敏感性分析处理不确定性因素。常用的敏感性分析方法主要有单因素敏感性分析、多因素敏感性分析、阙度法、极端分析法和概率分析法。在实际应用中,建议根据资料特点和分析目的采用多种敏感性分析方法进行分析。

## 7.6 定性研究

### 7.6.1 定义

定性研究是指在自然情境下采用多种资料互动对其行为和意义建构获得解释性理解的一种方法。从通俗意义上讲,定性研究能够回答一些定量研究和统计数据所不能回答的,诸如“为什么”、“怎么办”等问题。

### 7.6.2 研究设计和方法学要点

定性研究在目前的医疗卫生研究领域里主要有 3 种方法。

(1) 观察研究(observational study)。研究者从对事件、现象或人的行为观察中获得资料,最原始但是能够提供最直接的第一手资料。(2) 访谈研究(interview study)。研究者通过有目的地询问受访者并与其交谈获得资料的方法,主要分为结构化访谈和非结构化访谈。前者多用于社会调查,后者指在自然情境下的对话中进行自由交流。处在这两者中间的也是医学研究适合使用的半结构化访谈和深度访谈。在医学研究领域,访谈法最为常见。(3) 实物资料研究(documentary source)。对从各种渠道获得既成的资料与信息进行分析。

此外,定性研究的方法还有焦点组讨论法(focus group)、德尔非法、群体决策法(the normal group)、共识法(consensus methods)以及案例研究法(case study)等。较为典型的定性研究方法包括访谈法、焦点组讨论法、扎根式参与观察法、自传式民族志、基于受试者报告的症状量表、共识会议法、受试者报告结局等。

无论是哪种方法,都需要基于文献学习、专业背景知识等,通过共识、讨论等方式凝练出需要进行研究的观察提纲或者问题提纲。具体的凝练方式包括头脑风暴、小组讨论、专家咨询等。提出的观察或访谈提纲经过预访谈等方式确定其可行性和合理性,最终确立。基于确定的观察或访谈提纲,进行定性研究的资料收集,对所收集的资料进行分析,从而生成研究报告。

定性研究强调研究者与被研究者之间的和谐关系,注重在自然情景下收集资料,研究策略更具灵活性与变通性,体现个体特征而不强对总体进行推论,主要收集非量化的文字资料,资料分析方法采用归纳法。具体来说,其研究涉及灵活化,研究目的“求异化”,研究者的角色人性化,研究背景(场地)现场化,研究工具主观化,研究方法多样化,假说构建“扎根化”,资料分析“意义化”(理解化),抽样方法特殊化。定性研究采用独特的研究方式和方法,获得其他研究无法获得的研究资料和信息,如对深层的、特殊的现象进行深入细致的描述与分析,能了解研究对象复杂的、内在的心理、生活经验,即人的内心主观世界,人的言语行为所包含的意义和信息,以及无法量化的整体背景下的个体研究等。

因此,定性研究可以用来开拓未知领域并挖掘未知因素,无论是从广度还是深度上。适用于存在未知领域、不确定范式、概念、原因、深层元素以及新问题的学科。由于以上特点,其适用于中医药领域的研究。虽然当前定量研究似乎越来越占据主导地位,实际上定性研究在理解中医药的作用与结局上扮演重要角色。不同于仅

提供疗效均值的定量方法,定性研究回答了更多的细节问题,例如中医药疗效如何,适用于哪些条件,通过何种途径,针对何种人群,以及所谓“有效”的真正含义。定性研究可以评估干预方法的作用范围和类型,发现未曾预期或预料的结果,以及研究者从未考虑过的问题。在中医学研究领域,由于涵盖了哲学、文化、信念等元素,而且人文和主观因素在临床应用中能够发挥作用,是在进行研究的时候不可忽视的部分,利用定性研究来研究和探讨中医学,是非常值得尝试且具有重大意义的。

### 7.6.3 研究实施与资料收集

本节对定性研究中较常用的定性访谈和焦点组讨论分别进行介绍。

(1)定性访谈法。定性访谈是基于日常生活的专业性对话,其过程至少涉及两个人:采访者和受访者。研究者需要针对访谈过程,事先制定一个简单计划,并在访谈过程中执行以保证访谈内容的深度,即“半结构化深度访谈”。做好充分准备是开展定性研究访谈的先决条件。

在准备阶段,定性访谈大纲需要包括主题、关键词或问题,研究者需要结合受访者的情况,如知识种类和程度,还要考虑他们在访谈中的积极程度,在访谈之前,研究者需要对访谈的深度进行思考,并将这些信息体现在访谈指导手册中,这一点也非常重要。

被选定的访谈对象将获得一份简短的介绍信,包括研究目的、研究背景、研究流程、研究者资料、访谈焦点、被访者入选原因、访谈地点、访谈开始和结束时间、征求被访者同意与允许被访者随时退出研究的文字信息和诚信条款,并告知被访者研究及其结果。

访谈分为 3 个阶段:简介部分、主体部分和结尾部分。简介部分,双方首先相互接触并取得信任,关于一些背景信息简单重复。主体部分,研究者可以采用多种方法对受访者进行提问,具体分为引入问题及深入访谈问题。设置问题的目的在于获得受访者的反馈。在访谈的结尾部分,研究者有责任用受访者接受的方式结束访谈,并决定是否通过再次访谈进行随访,并了解受访者希望从研究者那里获得何种反馈。

(2)焦点组讨论法。焦点组即为了探讨一系列特定问题而聚集到一起的一组人群。一般一个典型的焦点组包括 8~12 名参与者和一名主持者,主持者鼓励参与者积极讨论和交流,其关注的重点是参与者之间的各种交流而不仅是对话。具体实施前需要进行的事项包括组的构建、选择和招募、参与者的选择和三角互证以及讨论主题提纲。具体操作包括主持人介绍研究主题、开场问题、关于研究主题的介绍性问题、讨论主体部分、结束讨论。

### 7.6.4 资料分析

(1)定性访谈的资料分析。获得了访谈资料后,需要通过几个步骤对资料进行分析:再听一遍访谈录音,评估音质并随时准备好发现新的研究问题。在决定如何誊录录音后,需要将所需部分转录成文字,再由其他人根据原始录音进行校对。接着进行转录内容的编码,此时研究者需要对所用概念有所选择,而此时分析步骤已经开始。用理论术语也可以进行编码,此步骤也可以通过计算机软件完成。

(2)焦点组讨论法。讨论之后,主持人和记录员需要互相询问关于焦点组讨论的执行情况,并进行资料分析,其中包括转录的语言以及语言外的社会元素。

## 7.7 系统综述

### 7.7.1 定义

系统综述和 Meta 分析属于循证医学系统综合的方法。系统

综述采用定性与定量相结合的方式,基于当前的临床信息,加以系统的综合和评价,从而为临床决策提供最佳证据。作为评价干预措施疗效的金标准,对所收集的研究资料进行定量的数据合并和综合分析,对所纳入的临床研究进行定性的方法学质量评价,系统地评估某种干预措施的疗效和安全性,从而为临床决策提供依据并且促进临床实践。

系统综述的定量分析中常见的统计方法是 Meta 分析。Meta 分析是采用统计学方法对两个或者两个以上的研究资料进行合并和综合的过程,一般情况下可以理解为系统综述的定量综合部分。在一项研究所纳入的临床试验中,如果两项或者两项以上的试验具有相同的干预措施、对照、同样标准的结局指标以及测量方法,或者研究间的一致性在合理的解释范围之内,则可以进行 Meta 分析将多项临床研究数据进行合并。值得注意的是,并非做了 Meta 分析的研究就是系统综述,因为后者必须是基于前瞻性检索策略和排除、纳入标准的制定以及系统的检索和资料收集,而前者只是一种统计学上的综合方法;相反,也并非所有的系统综述都有 Meta 分析,如果一项系统综述当中没有可以合并的临床研究,则可以不进行定量的数据分析,而仅仅进行描述性的文字综合以及定量的方法学质量评估。

### 7.7.2 系统综述的步骤和方法

系统综述在循证医学领域内,被视为鉴定并获取证据的最佳方法,一个系统综述通常纳入某一类研究类型的研究文献,最重要的就是纳入 RCT,尤其对于评价疗效的研究来说。Cochrane 协作组对 RCT 进行的系统综述被国际公认为高质量的系统综述,进行 Cochrane(或非 Cochrane)系统综述通常包含以下 7 个步骤。

(1)提出研究问题,形成研究方案。系统综述作为一种科研工作,研究方法需要预先确定。研究者在研究方案中需要考虑题目、研究背景、目的、纳入评价的研究标准、检索策略、评价方法、致谢、利益冲突、参考文献及附表。研究背景中应该提出需要解决的临床问题的合理性根据、重要性、意义及需要解决的途径。

由于系统综述是一项前瞻性的研究,目前在设计阶段进行研究方案的注册是越来越被接受和重视的步骤。Cochrane 系统综述在题目注册以后首先发表的就是研究的方案,研究者按照既定的格式撰写方案,经同行评议及编辑后发表。非 Cochrane 注册的系统综述题目,在撰写完成后可直接投稿给选定的期刊,但仍然提倡鼓励研究者事先注册方案。预先设计的系统综述可为医疗政策和临床决策的制定提供最好的证据。但是,研究者在进行系统综述和荟萃分析时,容易由于纳入研究的特征和方法学质量的评估结果与预想不同而修订预计的方案,使得系统综述的结论产生偏倚。同临床研究相似,对系统综述方案施行网络注册登记,可在一定程度上解决上述问题。研究者在进行相关综述前,需上传具体的方案,相关领域的评审专家和同行即可对其方案进行评论,提供相应的意见和建议,研究者可在此基础上对方案进行修正,然后按照方案进行系统综述的研究,在此基础上得出的结论会更有说服力。目前权威的注册机构是约克大学评价与传播中心(Centre for reviews and dissemination, CRD)建立的系统综述国际化前瞻性注册数据库(<http://www.crd.york.ac.uk/prospero/prospero.asp>, PROSPERO)。该平台为治疗、预防、诊断、监测等方面研究的系统综述研究者提供了免费的注册和审核流程,注册时需提供较详尽的计划书,按照网站提示必填的 22 个条目和选填的 18 个条目进行填写。一般提交方案后注册成功的时间在一周以内。注册后的方案会获得相应的注册编码,研究者在撰写系统综述全文

时可同时提供 PROSPERO 的注册编码,方便读者直接查询到相应的方案。

(2)研究的获得与选择。根据检索策略进行全面无偏的检索是系统综述与传统综述的关键区别。常用的数据库包括 PubMed、EMBASE、Cochrane 图书馆等英文数据库以及中国知网(CNKI)、维普、万方、SinoMed 等中文数据库。除了进行网络数据库的电子检索外,还包括手工检索发表或未发表的资料。通常检索不限制语言以防止语言偏倚。继而,根据研究方案中规定的符合系统综述的纳入标准和排除标准,对已经检索到的命中题录进行遴选。首先根据阅读题名和摘要进行一轮筛选,余下的题录下载全文后,阅读全文并进一步确定最终纳入系统综述的研究。为了准确客观,这项工作一般要求两名研究人员独立选择纳入的研究,出现不一致的情况由第三人或双方讨论协商解决。

(3)资料提取。主要涉及研究特征(如方法、对象、干预措施、结局)。方法通常包括设计类型、质量(如随机分配方案的产生、随机方案隐藏、盲法、病例退出情况、潜在的混杂因素等);研究对象包括种族、性别、年龄、诊断标准、研究背景、病例来源、纳入/排除标准等;干预措施包括试验和对照干预的名称、使用剂量与途径、实践、疗程以及有无随访等;结局测量可有多种结局,如病死率、发病率、生活质量、不良反应等,或同一结局采用不同的测量方法和测量时点。

(4)对纳入研究的质量进行评估。包括真实性和可能存在的各种偏倚。目前尚无质量评估的金标准方法,Cochrane 系统综述常用的质量评价标准为 Cochrane 协作网的 RevMan 软件中提供的偏倚风险(risk of bias, ROB)量表。

(5)分析与结果描述。根据系统综述所纳入的资料性质有定性和定量两种分析方法。定量的统计学分析又称 Meta 分析,而定性的分析主要指对单个研究进行描述性综合。

(6)结果解释(讨论)。主要涉及证据的强度、结果的可应用性、其他与决策有关的信息和临床实践的现状以及干预措施的利与弊、费用的权衡。

(7)系统综述的改进与更新。当有新的临床研究证据出现,就应当进行系统综述的更新,Cochrane 系统综述要求每年更新一次。

### 7.7.3 系统综述的资料分析

系统综述的目的是对收集的研究资料进行综合分析,确保结果的真实可靠,也就是要对某一干预措施的效果和(或)安全性进行全面评价,得到一个综合的结论,以指导决策或促进临床实践。在进行系统综述的过程中,要求从符合纳入标准的研究中提取原始资料,并用统计学方法对这些资料进行分析和概括。如果方法应用得当,数据的合成将为从这些资料中得出有意义的结论提供有力的帮助,同时也有助于避免在解释资料时发生错误。

系统综述的资料分析按照分析对象和分析方法不同,通常可分为定性的资料分析与定量的资料分析。

定性的资料分析是对单个研究的结果进行描述性综合,通常在各研究间存在明显异质性时进行。对资料的定性分析通常可包括对资料类型、相对效应、研究特征、研究结果等进行的描述性分析。

定量资料的统计学分析即 Meta 分析,在纳入研究干预、结局测量和测量方法、效应量的表达一致时可以进行。

首先,进行 Meta 分析需要选择效应量,常见的资料类型包括 3 种。一是计数资料,主要指二分类变量,如存活与死亡,有或无等。此类资料可用比值比(odds ratio, OR)、RR、相对危险度降低

(relative risk reduction, RRR)表示。第二种为连续变量,如身高、体重等,可用均数和标准差来表示,系统综述中常用组间均数的差值(means differences, MD)或标准化的均数差值(SMD)来合并效应量。第三种为生存率资料或时间-事件资料,常见于癌症的研究,这类资料通常用危害率(hazard ratios, HR)表示。

当研究间没有异质性,或其异质性在合理的解释范围且可用统计学方法予以处理,可采用固定效应模型进行 Meta 分析,而当所纳入的研究间存在较明显的临床异质性或者统计学异质性,则应采用随机效应模型,当临床异质性很大时,则需要放弃 Meta 分析,而仅进行描述性分析。当存在较明显的异质性时,在解释合并结果时应该格外小心,研究者应设法解释异质性的来源,如果能够在方案中事先设置好那些可能导致纳入研究结果出现差异的因素,将是更理想的。

由于研究间存在显著的异质性,可以考虑针对系统综述中的某些研究对象特征,如性别、年龄等进行亚组分析,而这也需要在研究方案制定时前瞻性作出计划。

发表偏倚是系统综述中存在的问题之一,它指的是阳性结果的研究容易得到发表的倾向。为了控制系统综述中由于发表偏倚导致研究结果偏离真实情况,首先,系统综述要求全面检索,因此对检索策略有严格的要求;其次,在纳入研究的时候,系统综述鼓励纳入未发表的研究,如会议摘要、学位论文等文献;同时,对于实际纳入的研究,用来检测发表偏倚的方法包括倒漏斗图、Egger 线性回归法、Begg 秩相关法、剪补法、失安全系数、Macaskill's 检验、Richy 法等。通常当纳入研究数量大于 10 时适宜采用倒漏斗图的方法,否则很难从图中对称性判断发表偏倚是否存在。

关于评价,推荐使用 Cochrane 协作网系统综述手册中风险偏倚评估工具[risk of bias (ROB) tool]对 6 个维度进行评价。

### 7.7.4 特殊类型的系统综述研究

(1)累积 Meta 分析。按原始研究发表的时间顺序进行的重复 Meta 分析,从而反映研究结果的动态变化趋势,评价各研究对综合结果的影响。

(2)个体患者数据(individual patient data, IPD)系统综述。与常规系统综述不同,IPD 系统综述需从原始研究作者手中收集纳入研究中每个受试对象的原始数据,将这些数据进行再分析。它获得的证据质量更高,被称为系统综述的金标准。

(3)前瞻性 Meta 分析。在纳入的任何研究结果尚未出来之前,先进行系统检索、评价和制定纳入排除标准的一种 Meta 分析方法。它可以体现“不确定性原则”,克服传统的回顾性系统综述存在的由临床试验结果已知而导致的偏倚。

(4)网络 Meta 分析。适用于临床试验中没有直接比较的证据,或者针对某一疾病/症状临床干预措施较多需要寻求最佳措施时,将传统的直接比较和间接比较同时合并起来进行 Meta 分析,形成网状形状。这一分析类型的主要作用就是对处于统一证据体的所有干预措施同时进行综合评价。

### 8 临床试验方案撰写与注册

临床试验方案是描述临床试验的背景、理论基础、目的、设计、方法和组织,包括统计学考虑、执行和完成条件的文件。它服务于多个项目:提供研究者指导试验实施的文件;提供受试对象描述详细的试验方法;提供伦理委员会事先了解受试者保障措施的文件以保护受试者的安全和权益;提供给资金支持方所提议的研究方法的评估手段;并且为系统综述作者以及其他读者提供评价潜在偏倚的描述。临床试验方案是涉及临床试验是否成功的关键文件,为开展的试验

项目的实施提供了依据、方法和标准,直接关系到是否能够按照试验目的和目标的要求对受试者进行干预,并客观准确地进行疗效和安全性观察或测量,而且是否能够在试验过程中遵守伦理学原则,为受试者提供安全、权益和福利保障。临床研究方案也是临床试验过程进行质量控制的重要参考依据。

8.1 临床试验方案的主要内容(表 3)

2007 年,115 个来自不同国家和地区的临床研究方案利益相关主体启动了规范临床研究方案内容,即“规范临床研究方案内容”(standard protocol items: recommendations for interventional trials, SPIRIT)。此国际合作项目通过循证医学为基础的推荐方法,提供临床研究方案的必要条目,从而提高临床研究方案的完整性。2013 年版本 SPIRIT 声明包含 33 个条目。

表 3 临床试验方案建议报告的条目以及相关文件

部分/条目	编号	描述
<b>管理信息</b>		
标题	1	题目应描述该研究的设计、人群、干预措施,如果适用也要列出题目的缩写
试验注册	2a	试验的标识符和注册名称。如果尚未注册,写明将注册机构的名称
	2b	WHO 临床试验注册数据所包括的所有数据集(附表,可查阅 <a href="http://www.annals.org">www.annals.org</a> )
方案版本	3	日期以及版本号
基金	4	基金的财政、物资和其他支持的来源和种类
角色与职责	5a	方案贡献者的名称、附属机构和角色
	5b	试验赞助者的名称和联系方式
	5c	如有试验资助者和赞助者,其在研究设计、收集、管理、分析及诠释资料、报告撰写、出版等环节的角色,以及拥有最终决策权者
	5d	试验协调中心、指导委员会、终点判定委员会、数据管理团队和其他监督试验的个人或团队的组成、作用及各自的职责,如果适用(参见 21 a 有关于数据监察委员会的内容)
<b>引言</b>		
背景及原理	6a	描述所要研究的问题以及开展试验的理由,包括对相关研究(已发表的与未发表的)中每个干预措施的有效性和不良反应的总结
	6b	对照组选择的解释
研究目的	7	特定的目的或者假设
试验设计	8	试验设计的描述,包括试验种类(如平行组、交叉、析因以及单一组),分配比例及研究框架(如优劣性、等效性、非劣势性、探索性)
<b>方法:受试对象、干预措施、结局指标</b>		
研究场所	9	描述研究场所(如社区诊所、大学的医院等)以及数据采集的国家列表,并且提供所列国家研究中心的网站以供参考
合格标准	10	受试对象的纳入、排除标准,如果适用需要说明研究中心以及实施干预的具体操作人员的合格标准(如手术、心理咨询师等特殊干预手段)
干预措施	11a	描述各组干预措施的细节以使他人能够重复,包括如何、在何时进行干预
	11b	中止或调整受试者所接受干预措施的标准(如针对不良反应、受试者要求、疾病进展或恶化情况等调整药物剂量)
	11c	提高依从干预措施方案的策略,以及任何监管方案遵守的流程(如药品回收、实验室检测等)
	11d	试验期间允许或禁止的相关伴随的护理和干预措施
结局指标	12	首要、次要和其他结局指标,包括具体的测量指标(如收缩压等),分析度量(如与基线的差值、终值、事件发生时间等),数据整合方法(如中位数、比例等),以及每个结局指标的测量时点。强烈建议对所选疗效和安全性结局指标的临床相关性做出解释
受试者时间轴	13	受试者纳入时间表、干预措施(包括所有磨合和洗脱情况),评价以及受试者访问。强烈建议使用示意图
样本含量	14	达到研究目的所需的受试者数目估计值,并说明如何确定样本含量,描述支持样本含量估算的临床假设和统计学假设
招募信息	15	描述为达到目标样本含量以纳入足够受试者的策略
<b>方法:干预措施分配(对照研究)</b>		
分配		
随机序列产生	16a	随机序列产生的方法(如计算机产生随机数字等),并列出所有分层因素。为降低随机序列的可预测性,应在单独的文件中记录任何限定的细节(如区组等),该文件对纳入受试者或分配干预措施的人员保密
随机分配隐藏	16b	用于执行随机分配序列的机制(如中心电话、按顺序编号不透光的密封的信封等),描述干预措施分配之前为随机序列隐藏所采取的步骤
实施	16c	谁生成随机分配序列,谁纳入受试者,谁分配受试者所接受的干预措施

续表 3

部分/条目	编号	描述
盲法	17a	分配干预措施后将谁设盲(如受试者、医护人员、结局评价者、数据分析人员等),以及如何实施盲法
	17b	如果实施盲法,试验期间在何种情况下允许破盲,描述暴露受试者所受干预措施的程序
方法:数据收集、管理与分析		
数据收集方法	18a	评价和收集结局、基线和其他试验数据的方案,包括任何提高数据质量的相关程序(如双测量、评价者培训等),并描述测量工具(如调查问卷、实验室检测等),如果测量工具的信度和效度已知,需进行描述,如果试验方案中没有提供数据收集表格,需提供参考文献
	18b	提供受试者完成治疗和随访的方案,包括列出中断或偏离干预措施方案受试者的结局数据
数据管理	19	数据录入、编码、安全和储存的方案,包括提供数据质量的相关流程(如双录入、数值范围检查等),如果试验方案中没有提供数据管理流程,需提供参考文献
统计方法	20a	分析首要和次要结局指标的统计学方法。如果试验方案中没有提供数据分析方法,需提供参考文献
	20b	其他分析的方法(如亚组分析和校正分析等)
	20c	不依从方案人群分析的定义(如随机化分析),以及处理缺失值的统计学方法(如多重填补等)
数据监察	21a	数据监察委员会的组成;委员会职责和监察报告的结构概述;申明是否独立,是否有利益冲突;如果试验方案中没有提供委员会章程的细节,需提供参考文献或说明不需要成立数据监察委员会的原因
	21b	描述中期分析和试验中止原则,包括谁可以使用中期结果和谁有权决定结束试验
危害	22	描述要求的和自愿报告的不良事件的收集、评价、报告和管理方案,以及干预措施或试验实施产生的其他意外效果
审计	23	如果有需报告审计试验实施的频率和流程,以及是否该过程独立于研究者和申办者
伦理与传播		
研究伦理批准	24	描述寻求伦理委员会批准通过试验伦理审查的方案
方案修正	25	描述与相关当事人(如研究者、伦理委员会、受试者、试验注册者、记录员和协调员等)交流重要的方案修正(如合格标准、结局指标、分析方法等的变化)
知情同意	26a	潜在的受试者中谁将会获得知情同意,以及如何实施知情同意(见条目 32)
	26b	若有辅助试验,需描述收集和使用受试者数据和生物学标本的其他知情同意条款
保密	27	描述如何收集、使用和保存潜在和纳入的受试者的个人信息,以保护在试验前、试验中和试验结束后的信息保密性
利益声明	28	整个试验的主要研究者和各个研究中心的财务和其他利益冲突声明
数据访问	29	申明谁有权使用试验最终数据集,并报告合同约定的对研究者该权利的限制
辅助和试验后护理	30	如果有辅助和试验后护理,描述其条款,并描述对那些因参加试验受到伤害的受试者进行补偿的条款
传播政策	31a	研究者和申办者将试验结果传达给受试者、医疗保健专家、社会公众以及其他相关组织的方案(如通过文献发表,在结果的数据中报告或其他分享数据的办法),包括出版限制条件
	31b	作者资格的指导原则以及拟使用的专业撰稿人
	31c	让公众可以查阅完整试验方案、受试者数据集以及统计学编码的方案(可选)
附件		
知情同意材料	32	知情同意书和其他相关的给予受试者和授权代理人的文件模板
生物标本	33	如果进行了辅助实验,描述收集、实验室评估和保存生物学标本,用作本试验和在未来辅助研究中进行基因或细胞分析使用的方案

## 8.2 临床试验的国际注册(表 4)

国际医学期刊编辑委员会(international committee of medical journal editors, ICMJE)要求所有临床试验在发表之前必须进行国际注册,否则不予发表试验结果。因此,对希望将研究成果发表在国际医学期刊上的国内医学研究人员来说,临床试验注册是必须加以重视的步骤之一。

目前符合世界卫生组织注册标准并获得 ICMJE 批准的试验注册机构和平台主要见下表。

注册试验的信息可直接提交给上述任何一个注册机构或中心,

但应按照不同国家对临床试验实施和监督过程的法律、伦理、资金、管理等方面的要求,尽可能在通过伦理审核、方案确定后,进行国际注册。其中美国临床试验注册中心(clinical trials.gov)在上述所有临床试验注册平台中,在临床试验注册数量和参与注册国家的广泛性方面居于首位。

## 9 统计分析计划(statistical analysis plan, SAP)

### 9.1 定义

SAP 是由生物统计专业人员根据试验方案要求起草的,比方案中描述的主要分析特征更具有技术性并提供更多信息的文件。它将

表 4 国际临床试验注册平台

名称	缩写	网址
Australian New Zealand Clinical Trials Registry	ANZCTR	<a href="http://www.anzctr.org.au/">http://www.anzctr.org.au/</a>
Brazilian Clinical Trials Registry	REBEC	<a href="http://www.ensaiosclinicos.gov.br/">http://www.ensaiosclinicos.gov.br/</a>
Chinese Clinical Trial Registry	Chi-CTR	<a href="http://www.chictr.org/cn/">http://www.chictr.org/cn/</a>
Clinical trials.gov	/	<a href="http://clinicaltrials.gov/">http://clinicaltrials.gov/</a>
Cuban Public Registry of Clinical Trials	RPCEC	<a href="http://registroclinico.sld.cu/">http://registroclinico.sld.cu/</a>
EU Clinical Trials Register	EU-CTR	<a href="https://www.clinicaltrialsregister.eu/">https://www.clinicaltrialsregister.eu/</a>
German Clinical Trials Register	DRKS	<a href="https://drks-neu.uniklinik-freiburg.de/drks_web/">https://drks-neu.uniklinik-freiburg.de/drks_web/</a>
International Standard Randomized Controlled Trial Number Register	ISRCTN	<a href="http://www.controlled-trials.com/">http://www.controlled-trials.com/</a>
Iranian Registry of Clinical Trials	IRCT	<a href="http://www.irct.ir/">http://www.irct.ir/</a>
Japan Primary Registries Network	JPRN	<a href="http://rctportal.niph.go.jp/">http://rctportal.niph.go.jp/</a>
Pan African Clinical Trial Registry	PACTR	<a href="http://www.pactr.org/">http://www.pactr.org/</a>
The Netherlands National Trial Register	NTR	<a href="http://www.trialregister.nl/">http://www.trialregister.nl/</a>
Sri Lanka Clinical Trials Registry	SLCTR	<a href="http://www.slctr.lk/">http://www.slctr.lk/</a>

分析数据集的选择、疗效和安全性的主要变量及次要变量、准备采用的统计分析方法、疗效和安全性评价方法、预期的统计分析图表等按统计分析报告格式拟定成计划书或模拟报告。从某种意义上来说, SAP 是从严格的统计学角度为回答试验方案中的研究目标或终点的临床问题而设计的, 为临床试验处理和分析数据的过程中坚持普遍接受的生物统计分析标准制定了基本规范。

## 9.2 SAP 的主要内容

SAP 形成于试验方案和病例报告表 (case report form, CRF) 确定之后, 是对统计分析报告的准备, 只涉及与统计分析相关的内容, 可根据实际问题在临床试验进行过程中进行修改、补充和完善, 在盲态审核时再次修改完善, 但是在第一次揭盲之前必须以文件形式予以确认, 此后不能再作变动。SAP 包括的首要内容是说明计划书中涉及的缩写, 还应包括以下内容。

### 9.2.1 简要说明临床试验的概要

相关内容包括: (1) 研究目的。研究目的是研究者希望通过本次临床试验解决的问题, 如评价某药物的疗效和安全性, 比较两种药物的经济学效益等。临床试验研究目的应该在试验方案制定过程中就确定。一个临床研究可以只有一个研究目的, 也可以有多个研究目的。对于多个研究目的的临床研究按其重要性分为主要目的和次要目的。(2) 研究的终点结局指标。包括主要结局指标和次要结局指标及探索性分析指标。(3) 说明临床试验的总体设计。试验设计、样本规模、研究药物、治疗计划、剂量调整与停药, 描述临床试验的总体设计原则, 受试者的入组方式, 样本规模, 研究药物的信息 (商品名、剂型、剂量、储存方式及生产企业等), 治疗方案 (治疗组和对对照组给予的药物、服药时间、服药方法), 剂量调整 (根据病程进行给药方案的调整), 停药的条件。

### 9.2.2 样本含量的计算方法及统计学把握度的设定

把握度是用数量描述的干预措施和对照措施之间如果确定有一个真正的差别存在, 能被显著性检验所检出的概率。研究设计时, 要求检出差别显著性的把握度越大, 其要求样本含量也越多。

### 9.2.3 确定统计分析方法

包括: (1) 根据试验目的确定比较类型, 如差异性检验、非劣效性检验、等效性检验、优效性检验。(2) 判断资料所属的设计方式, 如完全随机、配对设计、随机区组、析因设计、交叉设计、重复测量、观察性研究等设计类型。(3) 判断需要分析的资料是定量资料、无序分类资料还是有序分类资料。(4) 确定反应变量是单变量、双变

量还是多变量。(5) 判断影响因素是单因素还是多因素。(6) 判断资料是否符合拟采用的统计分析方法的应用条件, 比如是否服从正态分布、方差齐性等重要条件, 必要时应进行变量变换。

### 9.2.4 确定统计分析数据集

对统计分析的数据集进行划分: 全分析集 (full analysis set, FAS), 即根据符合 ITT 原则把所有随机化的受试者纳入分析的数据集, 是疗效评价的主要数据集。符合方案分析集 (per-protocol set, PPS), 即将符合方案规定完成药物治疗、无明显方案偏离、完成所有评价内容的受试者纳入分析, 它是全分析集的一个子集。方案偏离指不符合纳入标准、依从性差等。安全性分析集 (safety set, SS), 指所有经随机化分组, 有一次研究药物并进行了至少一次安全性评估的受试者构成的分析集, 用于评价药物安全性的主要人群。

规定缺失值和离群值的处理方法如下。缺失值是指数据收集过程中应该有的数据但未收集到。任何临床试验都会因为各种原因遇到数据缺失, 对于完全随机缺失, 可以忽视缺失值, 但是这种处理方法违背了 ITT 原则并可能对结果产生倚偏, 故不推荐为验证性临床试验的缺失值处理方法。其他缺失值处理方法包括简单填补 (末次访视截转、基线访视截转、最差病例填补、最好病例填补等), 多种填补 (多重回归填补法、趋势得分法、数据扩增法等)。离群值指收集的数据中存在的不合逻辑或常理的异常值。通过统计软件检测的异常值还需要医学专业知识的判断是否为离群值。

### 9.2.5 统计分析报告的呈现

统计分析报告一般以文字描述以及预设的表格和图等形式呈现, 包括: (1) 病例的分布和受试者的一般信息。(2) 治疗依从性。说明研究的依从性, 是否有脱落、失访病例以及原因。(3) 疗效评价。主要疗效指标表和次要疗效指标表。(4) 安全性评价。包括生命体征、实验室检查及心电图反应的情况, 不良事件和合并用药情况。

### 9.2.6 期中分析和数据监察委员会

该项内容应说明期中分析的目的、日程以及安排, 尽量避免计划外的期中分析。期中分析过程应保证能够控制整体的 I 类错误, 并由独立的数据监察委员会执行, 明确委员会的职责。

### 9.2.7 统计分析程序的设定

首先确定待分析的变量参数, 即描述主要结局指标和次要结局指标的定义及其评价标准和测量方法; 其次定义统计分析软件和结果输出模式, 一般数据集分析规则或定义等。

## 10 CRF

CRF 是根据研究方案设计的,用来在试验过程中记录每个研究对象的信息。CRF 作为临床研究中的二手数据,可由临床研究协调员抄录,经研究者审核后签字确认。临床数据交换标准协会 (clinical data interchange standards consortium, CDISC) 发布的临床数据采集标准 (clinical data acquisition standards harmonization, CDASH) 是目前业内公认和普遍采用的数据采集标准。CDASH 将数据采集字段按照其数据信息归属的类别分为一般标识变量、时间变量和 16 个领域,基本涵盖了临床研究中需要采集的数据信息。采用统一的数据标准将减少后期数据传输和交换过程中的工作量。

### 10.1 CRF 的设计流程

一般情况下,CRF 的设计由临床研究者负责,但期间数据管理人员应密切参与 CRF 的设计过程,以保证研究过程中搜集的数据满足其方案的要求。CRF 应符合设计规范,CRF 的封面不应出现受试者可识别的信息。

数据管理人员往往需要对定稿的 CRF 进行注释。注释的 CRF (annotated CRF, aCRF) 将作为一个重要文件进行存档。aCRF 中的变量名应与数据库中的变量名一致。而变量名的命名规则虽然在不同的机构有自身的惯例,但是采用 CDISC 体系标准进行注释,将节约后期数据传输和交换过程中的工作量。CRF 的设计一般经过形成初稿、修改与审查、批准与修订三个环节。

#### 10.1.1 形成 CRF 初稿

研究方案形成但尚未定稿时,CRF 设计者应着手 CRF 的设计,找出研究方案要求的与有效性和安全性指标相关的数据项,确定随访频次和时间窗,初步明确所要收集数据的频次、时间、内容、结构和类型,形成 CRF 初稿。提早进行 CRF 设计,可及时发现或确定研究方案中的一些细节问题,对其可能改进之处提出建议。如研究方案要求收集用药史,但未具体说明收集的时限范围;如方案要求收集不良事件,但实验室检查指标异常是否报不良事件、原有疾病加重是否报不良事件,复发性疾病如何记录不良事件等未进行详细说明;如要求记录方案偏离情况,但未定义方案偏离或定义不具有可操作性。总之,CRF 的设计是与方案的不断完善和优化同时进行的。

#### 10.1.2 CRF 的修改与审查

研究方案定稿后,检查 CRF 初稿中数据收集的内容、流程和格式是否符合研究方案的要求。对含义不明确的条目给出定义或解释,如专用术语、诊断与评分标准,注意检查对其解释是否与方案规定一致。

对 CRF 进行反复的审查和修改,审查人员通常为主要研究者、研究协调员 (clinical research coordinate, CRC)、项目经理、数据管理员、统计人员。审查的过程可帮助研究项目组成员进一步理解试验要求,并从各自的角度对 CRF 的内容和格式提出修改建议,使 CRF 满足各项目组成员的要求,最终对 CRF 的设计和使用达成共识。对 CRF 的审查,可制定 CRF 质量审查清单,要求项目组成员逐一审查,这样可避免遗漏审查项目,也便于全面收集各方意见。最后可进行模拟病历填写,以进一步验证 CRF 设计是否科学严谨,完整准确,简明美观,具有可操作性。

#### 10.1.3 CRF 的批准与修订

最后,依据申办方或研究基地的批准程序,完成 CRF 的批准后,可正式投入使用。在研究过程中,如修改研究方案,CRF 需要修订时,应严格按照变更控制程序进行修订,标示出相对于不同版本试验方案的 CRF 版本号,以免使用的 CRF 版本与最新试验方案

不符。

### 10.2 CRF 的主要内容

(1) 首页。一般包括试验名称、中心编号与名称、受试者编号、研究者姓名、临床试验方案编号等信息。(2) 填表注意事项。(3) 临床试验流程图。(4) 表头。一般包括临床试验批件号等信息。(5) 注脚。一般包括 CRF 版本号、版本日期等信息。(6) 基本试验参数模块。一般包括访视名称、访视日期、人口统计学信息、研究结束情况、研究者审核签名等。(7) 安全性相关模块。一般包括纳入排除标准、生命体征、体格检查、病史、伴随治疗、研究药物服用情况、实验室检查、不良事件等。(8) 疗效性相关模块。根据研究目的选择。(9) 预备与追踪 CRF 页。时点外访视所需模块。

### 10.3 CRF 设计的注意事项

(1) 文字叙述要准确,避免含糊其辞。(2) 格式要简洁易行,避免繁琐和复杂。(3) 用流程图表示必须进行的观察或检查的顺序。(4) 为待收集的数据留下足够的空白。(5) 只收集试验方案要求的数据,明确数值类型、数据位数与单位。(6) 采用问答形式,避免诱导性提问,填写内容尽量采用选项式,收集分类型数据,减少收集自由文本。(7) 避免重复收集数据。(8) 尽量采用业内公认的数据采集标准构建 CRF。(9) 提供 CRF 填写指南。

### 10.4 CRF 设计的常用软件

EpiData 软件是设计 CRF 的常用软件,该软件由丹麦的一个非盈利组织开发,是一个占用系统资源小,操作简便的免费软件。软件的录入界面与 CRF 的形式一致,方便数据录入,并且具有强大的数据核查功能,通过双录入控制数据录入的质量。

典型的 EpiData 数据库包含 3 个核心文件:调查表文件,用来定义 CRF 的结构;数据文件,用于存放数据以及已经定义好的编码;核查文件,用来定义数据输入时进行字段控制的有效性规则。

## 11 临床研究的伦理审核

### 11.1 临床研究伦理学应当参考的文件或规范

实施临床研究应当根据《赫尔辛基宣言》、国际医学科学组织委员会 (Council for International Organizations of Medical Sciences, CIOMS) 的《人体生物医学研究国际伦理指南》、卫生部《涉及人的生物医学研究伦理审查办法(试行)》、SFDA《药物临床试验质量管理规范》、《药物临床试验伦理审查工作指导原则》、国家中医药管理局《中医药临床研究伦理审查管理规范》等规范或文件结合临床实际情况进行试验方案的设计与实施。

### 11.2 伦理审核所需提交的材料

临床研究开始之前,首先应当确定相关的伦理审核机构,然后向其伦理委员会提交伦理审核所需的材料,主要包括:(1) 伦理审核文件的清单;(2) 伦理审核申请表;(3) 相关管理部门批准的批件;(4) 申办者资质证明;(5) 试验用品提供单位的资质证明;(6) 药品质量检测报告;(7) 申请项目的临床前整套研究资料摘要;包括综述资料、药学研究资料、药理毒理研究资料、对该项目迄今的临床经验总结,以及对照药品质量标准和临床研究文献资料;(8) 临床研究方案(版本日期和版本号);(9) 向受试者提供的研究简介和知情同意书(版本日期和版本号);(10) 用于招募受试者的材料(包括布告、广告等);(11) 受试者日记卡和其他问卷表(注明版本号和日期);(12) 主要研究者履历(最新版本,签名并注明日期);(13) 研究者手册(版本日期和版本号);(14) 研究 CRF(版本日期和版本号);(15) 如为多中心合作的试验,还需递交组长单位伦理委员会对申请项目的重要决定(包括否定结论或修改方案)的复印件和对研究方案作修改的说明;(16) 资金来源说明;(17) 不良反应应对措施

说明,文件中要对临床试验所带来不良反应的监测、出现不良反应时的应对措施及上报受试对象不良反应的渠道进行详细说明,可以作为方案中的应急预案;(18)临床研究项目还需要提交任务书或者立项书的复印件。

对初始审核未通过的研究项目,在申请复审时,除递交以上材料外,还应递交初始审核时伦理委员会对该研究项目的重要决定(包括否定结论或修改方案)及对研究方案做修改的说明。

## 12 数据管理

临床研究数据管理工作贯穿整个临床试验过程,其一般流程为:制定数据管理计划→设计数据采集工具(研究病历/CRF)→建立数据库→为临床研究人员提供数据管理培训→数据采集→数据核查与清理→盲态审核→数据库锁定→数据管理文件归档。研究数据收集过程中应能够保证信息得以保存,保留及恢复,且易于审核。对于数据相关的文件记录,“没有记录,等于没有做”(no document, no done)已经成为所有参与研究的人员必须遵守的一个重要原则。完整记录的文件体系,不仅可以体现研究人员已经完成的工作数量和质量,还可以证明研究的任何环节都遵循最好的和标准的操作规程(good practice)。临床研究数据管理学会(society for clinical data management, SCDM)发布的良好临床数据管理规范(good clinical data management practices, GCDMP),为临床研究数据管理工作提供了目前最全面、最详细的操作指南。药品审评中心(center for drug evaluation, CDE)于2012年发布了《临床试验数据管理工作技术指南》旨在加强对临床试验数据管理工作的监督管理。近年来,基于电子数据获取(electronic data capture, EDC)的数据管理发展迅速,电子化的数据管理具有提高试验效率和数据质量、节约成本、缩短研究周期等多方面优势,已逐渐成为数据管理的主流形式。

### 12.1 数据管理计划

数据管理计划(data management plan, DMP)是临床研究数据管理工作的纲领性文件,它对数据管理工作进行详细规划,确保及时获得真实、准确、完整、合法、可直接用于统计分析的数据库。DMP是临床研究的必备文档,临床试验启动前,数据管理员将根据研究方案、数据管理SOP及已有的监察计划、SAP、CRF来撰写DMP,一般经统计人员、临床监察员(clinical research associate, CRA)或项目经理审核批准后生效,作为临床研究有关各方开展日常工作的重要参考或依据。尽管DMP是预先规划的,是研究人员对未来研究数据操作流程的规划,但是在研究过程中,如果有重大事项的更改,DMP也须进行必要的修订。

不同机构的DMP在组织结构上可能有差异,以下13个方面基本涵盖了DMP的主要内容。

(1)批准页;(2)方案介绍;(3)工作职责与培训;(4)数据管理工作时间表;(5)数据库的设计、建立与维护;(6)数据录入;(7)数据核查与清理;(8)质量控制与质量保证;(9)外部数据的管理;(10)严重不良事件(serious adverse event, SAE)一致性核查;(11)医学编码;(12)数据库锁定;(13)数据管理报告与盲态审核。

### 12.2 数据库设计

糟糕的数据库设计方案不仅会严重影响数据的录入过程,还会影响后续的数据清理、提取和进行列表的程序的撰写和分析。数据库设计的主要目标是保证数据的准确录入,其他诸如以下的需求也应考虑:(1)清晰、便利和数据录入的速度;(2)能够有效地为统计

分析人员创建统计分析数据集;(3)数据格式的转换;(4)数据库应用的软件要求等。

### 12.3 数据录入

处理受试者信息数据时要做到保密且符合数据保护法,但是也要保证准确性。受试者治疗前的状况,以及对治疗的反应,包括评价者的观察,受试者的感觉以及不良反应,都应全面记录。处理受试者信息数据时要做到隐私保密且符合《保密法》、《数据保护法》。应尽一切努力保持记录无差错,并尽可能少丢失信息。试验参与者应得到足够的支持,以使他们能在试验过程中提供完整数据。使用网络这种现代收集信息的方法或许对准确记录数据有帮助。

### 12.4 数据清洗

在进行数据核查之前,数据管理员应制定数据核查计划(data validation plan, DVP),应对方案中规定的主要和次要有效性指标、关键的安全性指标进行充分的核查以确保数据的正确性和完整性。数据核查应在未知试验分组情况下进行,生成的数据质疑表(data clarification form, DCF)内容应避免有偏差或诱导性的提问。

数据的不一致(discrepancy)情况的发现可能来源于以下几个方面:(1)临床监察员或数据管理人员的手工检查;(2)由数据管理人员事先编写的计算机应用程序进行的实时的录入核查;(3)由数据管理人员或统计分析人员编写的独立于录入系统的外部程序核查。

### 12.5 盲态审核与数据库锁定

临床研究阶段完成数据的录入与清理后,由数据管理员报告试验过程中的数据管理情况,并由申办方、主要研究者、数据管理人员和统计师对临床研究数据进行审核和讨论,确定人群划分并定稿SAP。在以上所有工作经确认完成后,数据管理员将在获得锁定批准文件后锁定数据库,并将数据库移交统计方。

一般情况下,数据库锁定后发现数据仍然存在疑问的情况是很罕见的。所以,对于数据的解锁,应该有充分的理由,解锁的条件和流程必须遵循相应的SOP并详细记录。

在数据解锁和修改后,应该有清晰的质量控制流程保证数据的重新锁定。

### 12.6 数据管理文件归档

数据管理过程中将生成的大量工作文档,数据管理部门应妥善存档,以备稽查。由于各机构数据管理SOP的不同,保存的工作文档也存在差异,一般包括如下文档:(1)研究病历/CRF;(2)数据管理计划;(3)DVP;(4)数据库设计文档;(5)数据库测试记录;(6)数据库发布函;(7)用户培训记录及账户使用申请书;(8)空白/含数据的CRF;(9)数据库光盘;(10)数据库交接记录(交统计方/申办方);(11)各中心实验室理化指标参考值范围;(12)数据管理报告/盲态核查报告;(13)数据库锁定记录。

## 13 临床研究相关报告规范

### 13.1 RCT的报告规范(表5~9)

为了提高RCT的报告质量,一个由专家和编辑组成的工作组制定了RCT临床试验报告的统一标准(consolidated standards of reporting trials, CONSORT)声明。CONSORT声明于1996年首次发表,并于2010年更新,由25个条目组成,该声明促进了对RCT的严格评价和解释临床试验报告的统一标准。同时还包括非药物试验的CONSORT扩展。

表 5 CONSORT 2010 核对清单与非药物试验的 CONSORT 扩展  
(针刺试验 STRICTA2010, 以及拔罐和太极拳的相关干预措施的报告规范在 CONSORT 条目 5 上的扩展)

章节/主题	序号	CONSORT2010 声明:对照检查的条目	非药物试验的 CONSORT 扩展中增加的条目
文题和摘要	1a	文题能识别是随机临床试验	在摘要中描述试验措施,对照,医护提供者,试验中心和盲法状态
	1b	结构式摘要包括试验设计、方法、结果、结论;具体的指导建议参见摘要 CONSORT(CONSORT for Abstracts)	
引言			
背景和目的	2a	科学背景和对试验理由的解释	
	2b	具体的目的或假设	
方法			
试验设计	3a	描述试验设计(如平行设计、析因设计),包括受试者分配入各组的比例	
	3b	试验开始后对试验方法所作的重要改变(如合格受试者的人选标准),并说明原因	
受试者	4a	受试者合格标准	条件允许时,详述试验中心以及实施干预者的合格标准
	4b	资料收集的场所和地点	
干预措施	5	详细描述各组干预措施的细节以使他人能够重复,包括干预时间和方法	精确的描述试验措施和对照措施的细节
结局指标	6a	完整而确切说明预先设定的主要和次要结局指标,包括时间和方法	
	6b	试验开始后结局指标是否更改,并说明原因	
样本量	7a	如何确定样本量	条件允许时,论述是否及如何基于医护提供者或中心进行聚类分析的细节
	7b	必要时解释中期分析和试验中止的原则	
随机方法			
序列的产生	8a	产生随机分配序列的方法	条件允许时,描述医护提供者如何被分配到每一个试验组
	8b	随机方法的类型,任何限定的细节(如分区方法和各区组样本量)	
分配隐藏机制	9	用于执行随机分配序列的机制(如按序编码的封藏法),描述干预措施分配之前为隐藏序列号所采取的步骤	
实施	10	如何生成随机分配序列,执行招募干预措施分配者	
盲法	11a	如果实施盲法,分配干预措施后对谁设盲(如受试者、医护提供者、结局评估者),以及盲法如何实施	分组方法是否对联合干预实施者设盲,设盲的方法及描述干预措施的相似之处
	11b	如有必要,描述干预措施的相似之处	
统计学方法	12a	用于比较各组主要和次要结局指标	条件允许时,论述是否及如何基于医护提供者或中心进行聚类分析的细节
	12b	附加分析的方法,如亚组分析和校正分析	
结果			
受试者流程(极力推荐使用流程图)	13a	随机分配到各组的受试者例数,接受已分配治疗的例数,以及纳入主要结局分析的例数	每组中实施干预的医护提供者或中心的例数以及每一个医护提供者治疗的患者例数或在每一个试验中心中接受治疗的患者例数
	13b	随机分组后,各组脱落或被剔除的例数,并说明原因	
干预的实施			实施过程中描述试验措施和对照措施的细节
招募受试者	14a	招募期和随访时间,并说明具体日期	

续表 5

章节/主题	序号	CONSORT2010 声明:对照检查的条目	非药物试验的 CONSORT 扩展中增加的条目
	14b	试验中断或停止原因	
基线资料	15	以表格列出每一组受试者的基线数据,包括人口学资料和临床特征	允许时,描述每组中的医护提供者(病例数量、资历、专业技能等)和中心(数量)
纳入分析的例数	16	各组纳入每一种分析的受试者数目(分母),以及是否按最初的分组分析	
结局和估计值	17a	各组每一项主要和次要结局指标的结果,效应估计值及其精确性(如 95%可信区间)	
	17b	对于二分类结局,建议同时提供相对效应值和绝对效应值	
辅助分析	18	所做的其他分析结果,包括亚组分析和校正分析,区分预先设定的分析和新尝试的分析	
危害	19	各组出现的所有严重危害或意外效应[具体的指导建议参见危害 CONSORT(CONSORT for Harms)]	
讨论			
局限性	20	试验的局限性,报告潜在偏倚和不精确的原因,以及出现多种分析结果的原因(如果有这种情况)	
可推广性	21	试验结果被推广的可能性(外部真实性、适用性)	根据试验涉及的干预、对照、患者以及医护提供者和中心得出的试验结果的可推广性(外部真实性)
解释	22	与结果相对应的解释、权衡试验结果的利弊,并考虑其他相关的证据	还要考虑对照的选择,缺乏盲法或部分盲法,各组医护提供者或中心专业技能的不一致性
其他信息			
试验注册	23	临床试验注册号和注册机构名称	
试验方案	24	如果有,在哪里可以获得完整的试验方案	
资助	25	资助和其他支持的来源(如提供药品),资助者所起的作用	

注:\* 极力推荐结合“CONSORT 2010 说明与详述”阅读本声明,因为其对所涉及的全部条目作了详细阐述;推荐必要时阅读关于群组随机试验、非劣效性和等效性试验、草药干预,以及药物试验的各种 CONSORT 扩展版;详见 [www.consort-statement.org](http://www.consort-statement.org)

表 6 草药随机对照临床试验的报告:CONSORT 声明细则

章节/主题	序号	内容
文题和摘要	1	参加者如何被分配入组(例如:随机)标题和摘要中至少有一处标出该试验中所应用的草药产品的拉丁名,入药部位和剂型
引言		
背景	2	科学背景和原理解释包括简短说明进行此项试验的理由和使用该特定草药制品的依据,如果可行,请报告是否有关于此药物适应证的新的或传统的研究
方法		
受试者	3	参加者的入选标准、数据收集的场所和地点如果要检验的是传统适应证,那么就要对这种传统理论和观念进行描述,例如:参加者纳入标准应该反映出支持这一传统适应证的理论和观念
干预	4	详细描述每组的干预措施,包括给药时间和方法
草药产品名称	4a	每种草药成分的拉丁双语名、植物学权威名和科名,常用名;正确的商品名(例如:商标名称)或提取物名称(例如:EG b - 761),制造商名称;该药品在试验实施地是否经过认证(注册,登记)
草药产品的特征	4b	生产该药品或提取物所采用的植物部位;药品类型[生药(鲜或干),提取物];提取所用溶剂的类型和浓度(例如:80%酒精,100%水,90%甘油等);草药提取比例(例如2:1);生药材的鉴定方法(例如如何鉴定,鉴定人是谁)和批号,说明是否贮存了凭证标本(例如保留样品)及其贮存地和编号
给药方案和定量描述	4c	用药剂量、疗程及其依据;所有的定量草药产品(含生药和添加剂)的每单位剂量药物的质量、浓度等指标(适当时,可用范围来表示),添加剂材料,例如黏合剂、辅料和其他赋形剂(如17%麦芽糊精,3%二氧化硅/片),也需要在文中列出;标准化产品,必须列出活性/标志性成分的每单位药剂量

续表 6

章节/主题	序号	内容
定性检验	4d	(1) 产品的化学指纹及其检测方法(设备和化学参比标准品)和检测者(如实验室名称),是否贮存了产品样品(如保留样品)及贮存地;(2) 描述进行过的全部特殊检验/纯度测定(如重金属或其他污染物测定),报告去除了哪些物质及去除方法;(3) 标准化:被标准化的对象(如产品中哪种化学成分)和方法(如化学过程或生物/功能性活性测定)
安慰剂/对照组	4e	对照/安慰剂的说明
研究人员	4f	描述研究人员情况(如培训和实践经验)
目的	5	特定目的和假说
结局	6	清楚定义主要和次要结局指标,说明进行过的用以提高测量质量的方法(如多次观测和结局评价者培训),结局指标应反映干预措施和适应证的基础理论
样本量	7	样本量如何决定,解释所有的期间分析和终止条件
随机序列的产生	8	产生随机序列的方法,包括任何限制的细节描述(如区组、分层)
分配隐藏机制	9	执行随机分配序列的方法(如有编码的序列信封或中心电话),说明序列是否在干预措施分配结束之前处于隐藏状态
实施	10	说明分配序列制作人、受试者登记人、受试者分配人
盲法	11	对受试者、干预措施实施者和结局评估者是否使用盲法,如使用盲法,如何评价盲法的成功
统计学方法	12	用于比较组间主要结局的统计学方法;附加分析方法,如亚组分析和校正分析
结果		
受试者流程	13	推荐用流程图报告各阶段受试者流程;特别是报告参加随机分组、接受治疗、完成研究方案、参加主要结果分析的受试者数目;描述实际研究情况与研究方案之间变异的情况及其原因
募集受试者	14	明确定义募集受试者时间和随访时间
基线资料	15	基线人口统计学和临床特征包括联合使用的医疗措施、草药和替代治疗
数据分析	16	纳入每一分析的受试者数(分母),是否采用了 ITT;如可能,采用绝对数字表述结果(如 10/20 而不是 50%)
结局和效应值	17	对每一个主要和次要结局说明每组汇总的结果、效应估计值及其精确性(如 95% 可信区间)
辅助分析	18	报告所进行的其他任何分析以说明方法的多样性,包括亚组分析、校正分析,指出预先制定和临时添加的分析
不良事件	19	各组所有重要不良事件或副作用
讨论		
结果解释	20	结果解释应考虑研究假设、潜在偏倚和不精确的原因,及与结果和分析的多样性相关的危险因素,根据产品/给药方案解释结果
可推广性	21	试验结果和结论的可推广性(外部真实性)可能时,讨论本试验所用的草药产品和给药方案与在自我保健和(或)临床实践中应用的关系
解释	22	根据当前证据,概括解释结果联系其他产品的试验,讨论本试验结果

表 7 针刺临床试验干预措施报告标准核对清单

条目	序号	内容
针刺治疗的合理性	1a	针刺治疗的类型(如中医针刺、日本汉方医学针刺、韩国韩医针刺、西医针刺、五行针刺、耳针等)
	1b	提供针刺治疗的理由,依据的历史背景,文化因素,和(或)共识法,引用文献(适当情况下)
	1c	不同治疗程度的详细介绍
针刺的细节	2a	每一受试对象每一治疗环节用针的数目(相关时用均数和变异范围表示)
	2b	使用的穴位名称(单侧/双侧)或位置(如无标准穴位名称则说明)
	2c	进针的深度,基于指定的计量单位,或描述进针的肌肉层次
	2d	引发的机体反应(如得气或肌肉抽动反应)
	2e	针刺刺激方式(如手工行针刺刺激和电刺激)
	2f	留针时间
	2g	针具类型(直径、长度和生产厂家或材质)

续表 7

条目	序号	内容
治疗方案	3a	治疗次数
	3b	治疗频率和持续时间
辅助干预措施	4a	对针刺组施加的其他附加干预措施细节(如灸、拔罐、中药、锻炼、生活方式建议)
	4b	治疗场所和环境,包括对治疗师的操作指南所作的说明和指导以及对患者所作的解释说明
治疗师的资历	5	对针刺治疗师的描述(资历或从业部门,从业时间,其他相关经历)
对照和干预	6a	基于研究问题解释试验中对照或对比干预的选择,援引资料证明其合理性
	6b	精确地描述对照或对比干预措施,如果采用假针刺或其他任何一种类似针刺的对照措施,需提供条目 1~3 所要求的细节信息

国际草药 CONSORT 声明沿用了 CONSORT 声明的 22 个项目,细化了其中的 9 项,以使其更加适应草药临床试验的报告,包括对 8 个项目的小的建议和对 1 个项目的细化建议[第 4 项(干预)]。

《针刺临床试验干预措施报告的标准》(Standards for Reporting Interventions in Clinical Trials of Acupuncture, STRICTA) 报告指南最早在 2001 年出版,其设计是为了提高针刺对照试验中干预措施报告的完整性和透明度,使这种试验能够更加清楚地被解释和重复。STRICTA 包含一个扩大了 CONSORT 声明中条目 4 的通用内容的清单,此清单与干预措施的报告有关。修订的 STRICTA 清单包括 6 个条目,17 个亚条目。

除针刺临床试验外,目前尚无关于太极拳及拔罐的临床试验干预措施报告规范建议,根据 STRICTA、CONSORT 声明及其非药物治疗扩展版制定出的关于太极拳及拔罐临床试验干预措施的报告规范建议条目,供各位读者参考。

13.2 观察性研究的报告规范(表 10)

2007 年,WHO 推荐为改善观察性研究论文的质量而建立的报告

规范——加强观察性流行病学研究报告的质量(strengthening the reporting of observational studies in epidemiology, STROBE),由于观察性流行病学研究常包含数种研究设计和诸多的主题,因此制定小组把 STROBE 建议限定在 3 种研究设计(队列研究、病例对照研究和横断面研究),并制定成通用的格式。共 22 个条目。

13.3 定性研究的报告规范(表 11)

目前暂时没有公认的权威的定性研究报告规范,但相比之下报告定性研究综合标准(consolidated criteria for reporting qualitative research, COREQ)更为具体和容易掌握,共 32 项。

13.4 系统综述的报告规范(表 12)

PRISMA 声明(preferred reporting items for systematic reviews and meta-analyses)由 27 个条目组成。本声明的目的在于帮助作者改进系统综述和 Meta 分析的撰写和报告。主要针对 RCT 的系统综述,也可作为其他类型研究系统综述报告的基础规范,尤其是对干预措施进行评价的研究。PRISMA 也可以用于已发表系统综述的严格评价。

表 8 太极拳临床试验干预措施报告建议条目清单

条目	序号	内容
太极拳干预的合理性	1a	太极拳的类型和套路(例如:108 式杨式太极拳、24 式简化杨式太极拳等)
	1b	所提供太极拳干预的理由、依据的历史背景、文化背景,和(或)共识法,在适当处提供参考文献
	1c	说明太极拳干预的在哪些方面发生变化
太极拳学习和练习的细节	2a	太极拳学习方法(例如:在太极拳老师的指导下学习、跟随太极拳视频或书籍自学等)
	2b	太极拳学习步骤(例如:首先进行准备活动,然后进行太极拳练习,最后进行整理运动等)
	2c	太极拳的强度和衡量标准
	2d	太极拳练习方法(例如:集体练习或自己单独练习;在太极拳老师的监督或带领下练习等)
	2e	太极拳学习内容(例如:太极拳理念、原则、呼吸技巧、太极拳动作等)
	2f	太极拳学习和练习过程(例如:试验前 2 周循序渐进地学习太极拳,后 8 周练习全套太极拳等)
太极拳干预方案	3a	太极拳干预的单元数
	3b	太极拳干预单元的频率和疗程
其他干预措施	4a	对太极拳组施加的其他干预措施的细节(例如:常规西药、中药、针刺、慢跑、自行车、生活方式建议等)
	4b	干预实施的场所和相关信息,包括对受试者解说的信息。
太极拳老师的资历	5	对太极拳老师的背景进行描述(包括资历、从属的太极拳相关机构、习练太极拳的时间,太极拳教学经验等)
练习太极拳的达标要求	6	针对该试验的研究目的和采用的太极拳干预措施,具体报告受试者习练太极拳的达标评估标准以及评价者
对照或者对照干预	7a	在阐述研究问题时引用资料证明选择对照或对照措施的合理性
	7b	精确地描述对照或对照措施,如果采用其他任何与太极拳类似的对照措施,则提供条目 1 到 3 所要求的细节信息

注:阅读该清单时,建议与每个条目的详细介绍和高质量报告范例一起阅读,本清单适合与 CONSORT2010 条目 5 关于临床试验干预措施报告的条目结合进行报告;表 9 同

表 9 拔罐疗法临床试验干预措施报告的信息项目核对清单

条目	序号	内容
拔罐的基本原理	1a	拔罐的种类及选择拔罐治疗的依据(如辨证情况、罐疗种类选取、个体化治疗选择等)
	1b	文献理论依据
	1c	不同治疗程度的详细介绍
拔罐的细节	2a	每一受试对象每一治疗环节用罐的数目(相关时用均数和变异范围表示)
	2b	穴位或行罐部位的选择(单侧/双侧)
	2c	行罐的手法及作用强度
	2d	不同拔罐方法行罐的时间等
	2e	对刺络拔罐研究应同时详细描述刺络的方法和穴位选择
	2f	对出血量的控制情况
	2g	罐的类型(生产厂家或材质)
治疗方案	3a	治疗次数
	3b	治疗频率和持续时间
辅助干预措施	4a	对拔罐组施加的其他附加干预措施细节(如针刺、灸、中药、锻炼、生活方式建议)
	4b	治疗场所和环境,包括对治疗师的操作指南所作的说明、指导以及对患者所作的解释说明
治疗师的资历	5	治疗师是否具有相关培训和资质证明、临床实际工作的年限以及对该疾病的治疗水平
对照和干预	6a	基于研究问题解释试验中对照或对比干预的选择,援引资料证明其合理性
	6b	精确地描述对照或对比干预措施,如果采用其他任何与拔罐相似的对照措施,需提供条目 1~3 所要求的细节信息

表 10 STROBE 声明——观察性研究必需项目清单

内容与主题	序号	内容
标题和摘要	1a	题目或摘要中要有常用专业术语表述研究设计
	2b	摘要内容要丰富,并且能准确流畅地表述研究中做了什么
前言		
背景/原理	2	对所报告的研究背景和原理进行解释
目标	3	阐明研究目标,包括任何预先确定的假设发现了什么
方法		
研究设计	4	在论文中较早陈述研究设计的要素
研究现场	5	描述研究现场、具体场所和相关时间范围,包括研究对象征集、暴露、随访和数据收集时间
研究对象	6a	队列研究:描述选择研究对象的合格标准、源人群和选择方法,描述随访方法;病例对照研究:描述选择确诊病例和对照的合格标准、源人群和选择方法,描述选择病例和对照的原理;横断面研究:描述选择研究对象的合格标准、源人群和选择方法
	1b	队列研究 - 配对研究:描述配对标准和暴露与非暴露数目;病例对照研究 - 配对研究:描述配对标准和每个病例对应的对照数目
研究变量	7	明确定义结局、暴露、预测因子、潜在的混杂因子和效应修饰因子,如果可能,列出诊断标准
数据来源/测量	8*	对每个关心的变量,描述其数据来源和详细的判定(测量)方法,如果有多组,还应描述各组之间判定方法的可比性
偏倚	9	描述和解释潜在偏倚的过程
样本大小	10	解释样本量的确定方法
计量变量	11	解释分析中如何处理计量变量,如果可能,描述怎样选择分组及分组原因
统计学方法	12a	描述所有统计学方法,包括控制混杂方法
	12b	描述亚组和交互作用检查方法
	12c	描述缺失值处理方法
	12d	队列研究:如果可能,解释失访的处理方法;病例对照研究:如果可能,解释病例和对照的匹配方法;横断面研究:如果可能,描述根据抽样策略确定的统计方法
	12e	描述敏感性分析
结果		
研究对象	13a	报告研究的各个阶段研究对象的数量,如可能合格的数量、被检验是否合格的数量、证实合格的数量、纳入研究的数量、完成随访的数量和分析的数量
	13b	描述各个阶段研究对象未能参与的原因
	13c	考虑使用流程图
描述性资料	14a*	描述研究对象的特征(如人口学、临床和社会特征)以及关于暴露和潜在混杂因子的信息
	14b*	指出每个关心的变量有缺失值的研究对象数目
	14c*	队列研究:总结随访时间(如平均时间及总时间)
结局资料	15*	队列研究:报告发生结局事件的数量或根据时间总结发生结局事件的数量;病例对照研究:报告各个暴露类别的数量或暴露的综合指标;横断面研究:报告结局事件的数量或总结暴露的测量结果

续表 10

内容与主题	序号	内容
主要结果	16a*	列出未校正的和校正混杂因子的关联强度估计值和精确度(如 95% CI), 阐明根据哪些混杂因子进行调整以及选择这些因子的原因
	16b*	当对连续性变量分组时报告分组界值
	16c*	如果有关联, 可将有意义时期内的相对危险度转换成绝对危险度
其他分析	17	报告进行的其他分析, 如亚组和交互作用分析及敏感性分析
讨论		
重要结果	18	概括与研究假设有关的重要结果
局限性	19	结合潜在偏倚和不精确的来源, 讨论研究的局限性; 讨论潜在偏倚的方向和大小
解释	20	结合研究目的、局限性、多因素分析、类似研究结果和其他相关证据, 谨慎作整体的结果解释
可推广性	21	讨论研究结果的可推广性(外推有效性)
其他信息		
资助	22	列出当前研究的资助来源和资助者, 如果可能, 列出原始研究的资助情况

注: \* 在病例对照研究中分别列出病例和对照的信息; 如果可能, 在队列研究和横断面研究里给出暴露组和未暴露组的信息

表 11 COREQ 项目清单

内容与主题	序号	内容
研究团队和过程反映		
研究者个人特征		
访谈者/组织者	1	哪位(些)文章作者实施访谈或焦点组访谈
学位/学历	2	研究者的学位, 例如: 理学博士(PhD)或医学博士(MD)
职业	3	在研究进行时, 研究者的职业
性别	4	研究者性别
经验和培训	5	研究者的经验和培训情况
研究者与参与者的关系		
关系建立	6	与参与者的关系是在开始研究前就已建立
参与者对访谈者的了解	7	参与者了解的访谈者信息, 如个人目标及研究依据和理由
访谈者特征	8	文中报告的访谈者/组织者特征, 如偏倚、研究结果猜测、进行研究的原因和兴趣
研究设计		
理论框架		
方法学观念和理论	9	文章报告的在研究中被应用的方法学观念、理论和方法, 如扎根理论、话语分析、人种学和内容分析
选择参与者		
抽样	10	如何选择参与者, 如目的性抽样、便利性抽样、连续性抽样及滚雪球抽样
与参与者沟通的方法	11	如何与参与者沟通, 如面对面、电话、信件或电子邮件
样本量	12	研究中有多少参与者
拒绝参加研究或中途脱落	13	多少例拒绝参加研究或中途脱落, 原因何在
场所		
资料收集场所	14	在哪里收集的资料, 如家里、诊所或工作场所
在场的非参与者	15	除了参与者与访谈者外, 是否还有其他人在场
样本描述	16	样本的主要特征, 如人口学信息和日期
收集资料		
访谈提纲	17	访谈中所用到的问题、提示和提纲等是否由文章作者提供, 是否经过预访谈检验
重复访谈	18	是否进行过重复访谈, 如果进行过, 有多少次
音/像录制	19	研究是否通过录音或录像收集资料
场记	20	在个体访谈/焦点组访谈过程中和(或)结束后是否作了场记
时长	21	个体访谈或焦点组访谈的时长
信息饱和	22	是否讨论了信息饱和问题
转录文字返还	23	访谈转录成文字后是否返还给参与者征询意见和(或)纠正错误
分析和结果		
分析资料		
资料编码的数量	24	共用了多少个代码对资料进行编码
描述编码树	25	作者是否描述了编码树
主题来源	26	主题是预设的, 还是源自获得的资料

续表 11

内容与主题	序号	内容
软件	27	如果用了软件来管理资料,软件的名称和必要信息
参与者检查	28	参与者是否提供了对研究结果的反馈
报告		
报告引文	29	是否用了参与者引文来说明主题或结果,每条引文是否都有身份标记,如参与者编号
资料 and 结果的一致性	30	根据报告的资料能否得出研究的结果
重要主题的清晰报告	31	研究结果中是否清晰报告了重要主题
次要主题的清晰报告	32	是否有对特殊案例的描述和对次要主题的讨论

表 12 系统综述或 Meta 分析报告条目清单

项目	序号	内容
标题	1	明确本研究报告是针对系统综述、Meta 分析,还是两者兼有
摘要		
结构式摘要	2	提供结构式摘要包括背景;目的、资料来源、纳入研究的标准、研究对象和干预措施、研究评价和综合的方法、结果、局限性、结论和主要发现、系统综述的注册号
前言		
理论基础	3	介绍当前已知的研究理论基础
目的	4	通过对研究对象、干预措施、对照措施、结局指标和研究类型五个方面 (participants, interventions, comparisons, outcomes, study design, PICOS) 为导向的问题提出所需要解决的清晰明确的研究问题
方法		
方案和注册	5	已有研究方案,则说明方案内容并提供可获得该方案的途径(如网址),提供现有的已注册的研究信息,包括注册编号
纳入标准	6	将指定的研究特征(如 PICOS, 随访的期限)和报告的特征(如检索年限、语种、发表情况)作为纳入研究的标准,并提供合理的说明
信息来源	7	针对每次检索及最终检索的结果描述所有文献信息的来源(如资料库文献,与研究作者联系获取相应的文献)
检索	8	至少说明一个资料库的检索方法,包含所有的检索策略的使用,使得检索结果可以重现
研究选择	9	说明纳入研究被选择的过程,包括初筛、合格性鉴定及纳入系统综述等步骤,还可包括纳入 Meta 分析的过程)
资料提取	10	描述资料提取的方法(例如预提取表格、独立提取、重复提取)以及任何向报告作者获取或确认资料的过程
资料条目	11	列出并说明所有资料相关的条目(如 PICOS, 资金来源),以及作出的任何推断和简化形式
单个研究存在的偏倚	12	描述用于评价单个研究偏倚的方法(包括该方法是否用于研究或结局水平),以及在资料综合中该信息如何被利用
概括效应指标	13	说明主要的综合结局指标(如危险度比值、均值差)
结果综合	14	描述结果综合的方法,如果进行了 Meta 分析,则说明异质性检验的方法
研究偏倚	15	详细地评估可能影响数据综合结果的可能存在的偏倚(如发表偏倚、研究中的选择性报告偏倚)
其他分析	16	对于研究中其他的分析方法进行描述(如敏感性分析或亚组分析、Meta 回归分析),并说明哪些分析是预先制定的
结果		
研究选择	17	报告初筛的文献数、评价符合纳入的文献数,以及最终纳入研究的文献数,同时说明每一步排除文献的原因,最好提供流程图
研究特征	18	说明每一个被提取资料的文献的特征(如样本含量、PICOS、随访时间)并提供引文出处
研究内部偏倚风险	19	说明每个研究中可能存在偏倚的相关数据,如果条件允许,还需要说明结局测量水平的评估(见条目 12)

续表 12

项目	序号	内容
单个研究的结果	20	针对所有结局指标(有效或有害性),说明每项研究的各干预组结果的简单合并以及综合效应值及其可信区间,最好以森林图形式报告
结果的综合	21	说明每个 Meta 分析的结果,包括可信区间和异质性检验的结果
研究间偏倚	22	说明对研究间可能存在偏倚的评价结果(见条目 15)
其他分析	23	如果有,列出其他分析的结果(如敏感性分析或亚组分析,即 Meta 回归分析,见条目 16)
讨论		
证据总结	24	总结研究的主要发现,包括每一个主要结局的证据强度;分析它们与主要利益集团的关联性(如医疗保健的提供者、使用者及政策决策者)
局限性	25	探讨单个研究和结局水平的局限性(如偏倚的风险),以及系统综述的局限性(如检索不全面,报告偏倚等)
结论	26	列出对结果的概要性的解析,并提出对未来研究的提示
资金支持		
资金	27	描述本系统综述的资金来源和其他支持(如提供资料),以及系统综述的资助者

**利益冲突声明**

本指南的制定与内容不存在利益关系。指南起草人与论证专家与药物企业无利益关系。

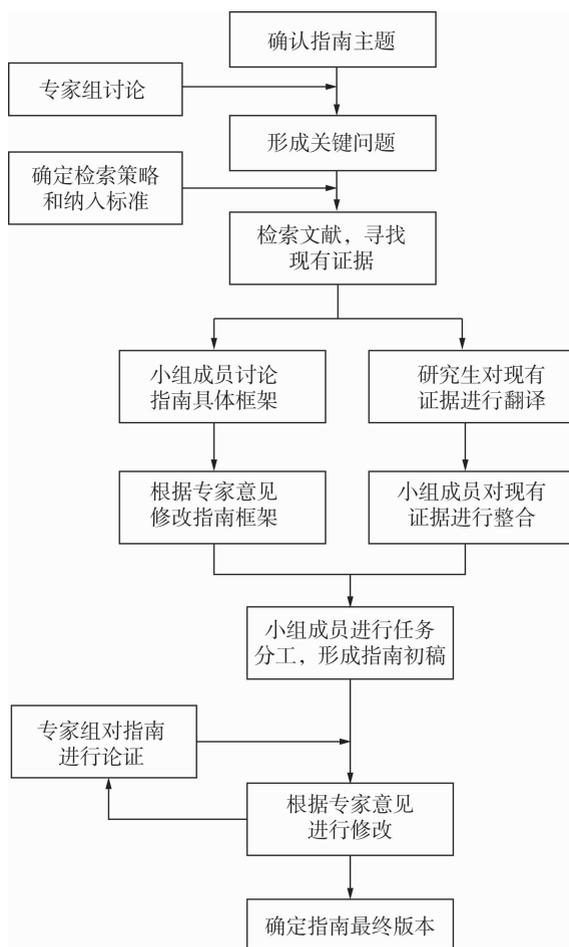
**致谢**

参与本指南制定的国际专家受到欧洲委员会 FP7 签署的 GP-TCM 项目的资助,项目资助协议号为 223154。Andrew Flower 医生目前为英国国家健康研究所资助研究员。George Lewith 教授受 Rufford Maurice Laing 基金会资助。北京中医药大学循证医学中心刘建平平等受美国国立卫生研究院项目 R24 AT001293 资助。本项目工作得到了国家中医药管理局行业专项“中医临床研究能力模式研究”(No.201207007)以及北京市重点学科共建项目(中西医结合循证医学)支持。循证医学中心博士研究生苏春香、周芬、陈妮妮、薛培、冯硕,硕士研究生王丽琼、杨国彦、任君、孙瑾、张抗、梁宁、李文元等参与英文的翻译工作。

**指南制作流程**

本指南制作基于循证医学和临床流行病学的基本原理和方法,系统查阅国际上适用于中医药临床研究最新方法学研究进展,结合著者 20 余年的教学和培训经验,编写而成。指南编写小组自 2014 年 5 月正式启动编写工作,部分内容采用了欧盟资助项目(GP-TCM)中第六工作组——《中医临床研究方法学组》成员所编写并发表的中药临床试验方法指南,该指南 2012 年摘录发表于 Journal of Ethnopharmacology,2014 年全文发表于欧洲整合医学杂志。经过编写小组委托研究生翻译,编写组核心成员进行审校,部分内容得以整合到本指南当中。指南初稿形成于 2014 年 11 月,经过编写组多次讨论、修改,自 2015 年 3 次征询中国大陆及中国香港地区相关临床研究方法学专家、中医药专家、中西医结合专家、临床流行病学与医学统计学专家、中医药科研管理专家等的意见,反复进行了修改、编辑和完善。

该指南将通过中国医师协会中西医结合医师分会和中国中西医结合学会循证医学专业委员会联合发布。指南发布寄希望于中医药临床研究在设计、实施和报告等方面能够更加规范,质量得以提高。并且通过该指南的推广和使用,提出新的问题和需求,欢迎任何反馈意见和建议,这将对编写组未来在指南更新和修订时提供重要的参考依据。



指南制作流程

本指南起草小组成员(按实际贡献大小排序):刘建平(北京中医药大学),陈可冀(中国中医科学院西苑医院),费宇彤(北京中医药大学),高蕊(中国中医科学院西苑医院),陈薇(北京中医药大学),曹卉娟(北京中医药大学),张颖(北京中医药大学),刘兆兰(北

京中医药大学),李迅(北京中医药大学),韩梅(北京中医药大学),王丽琼(北京中医药大学),王梅(辽宁中医药大学),杨国彦(北京中医药大学),于河(北京中医药大学),夏芸(北京中医药大学东方医院),陈妮妮(北京中医药大学),冯硕(北京中医药大学),孙瑾(北京中医药大学), Andrew Flower (英国南安普顿大学), Claudia Witt (德国柏林夏瑞塔大学), George Lewith (英国南安普顿大学), Gudrun Ulrich-Merzenich (德国柏林大学), Kenneth Muire (英国华威大学), Martina Prude (英国南安普顿大学)。

本指南发起人:陈可冀(中国科学院院士)、刘建平教授(长江学者)

本指南论证专家(按姓氏笔画排序):王文健(复旦大学附属华山医院)、王辰(卫生部中日友好医院)、王泓午(天津中医药大学)、王思成(国家中医药管理局科技司)、卞兆祥(香港浸会大学中医药学院)、方敬爱(山西医科大学第一医院)、吕爱平(香港浸会大学中医药学院)、刘鸣(四川大学华西医院)、刘建平(北京中医药大学循证医学中心)、时景璞(中国医科大学第一医院)、吴大嵘(广州中医药大学第二附属医院)、吴盛贤(北京中医药大学东直门医院)、张伯礼(中国中医科学院)、张博恒(复旦大学附属中山医院)、陈可冀(中国中医科学院西苑医院)、陈香美(中国人民解放军总医院)、赵一鸣(北医三院临床流行病学研究中心)、赵亚双(哈尔滨医科大学)、徐浩(中国中医科学院西苑医院)、翁维良(中国中医科学院西苑医院)、高蕊(中国中医科学院西苑医院)、唐旭东(中国中医科学院西苑医院)、唐金陵(香港中文大学公共卫生学院及基层医疗学院)、黄光英(华中科技大学附属同济医院)、康德英(四川大学华西医院)、梁伟雄(广州中医药大学附属第一医院)、谢雁鸣(中国中医科学院临床基础研究所)、詹思延(北京大学公共卫生学院流行病学与统计学系)、蔡定芳(复旦大学附属中山医院)、熊宁宁(南京中医

药大学江苏省中医院)。

指南工作小组:王丽琼、陈惠、刘龙涛、付长庚、高蕊、陆芳、张艳萍、李顺华

支持单位:《医师报》社

### 参 考 文 献

- [1] Witt CM, Aickin M, Cherkin D, et al. Effectiveness guidance document (EGD) for Chinese medicine trials: a consensus document [J]. *Trials*, 2014, 15:169.
- [2] Flower A, Witt C, Liu JP, et al. GP-TCM Unbridged guidelines for randomized controlled trials investigating Chinese herbal medicine (CHM) [J]. *Eur J Integr Med*, 2014, 6(2): 186-210.
- [3] Flower A, Witt C, Liu JP, et al. Guidelines for randomized controlled trials investigating Chinese herbal medicine [J]. *J Ethnopharmacol*, 2012, 140(3): 550-554.
- [4] 黄悦勤.临床流行病学.第4版[M].北京:人民卫生出版社,2014.
- [5] 刘建平.循证医学[M].北京:人民卫生出版社,2012.
- [6] 李立明.临床流行病学[M].北京:人民卫生出版社,2011.
- [7] 刘建平.循证中医药定性研究方法学[M].北京:人民卫生出版社,2009.
- [8] 刘建平.循证中医药临床研究方法[M].北京:人民卫生出版社,2009.

## 《中国中西医结合杂志》荣获中国科协精品科技期刊项目

按照《中国科协精品科技期刊工程项目实施方案(2015—2017)》,通过项目申报,资格审查及专家评审,分别评选出精品科技期刊 TOP50 项目 50 项,学术质量提升项目 120 项,数字出版建设项目 11 项,集群(联盟)建设项目 11 项,出版人才培养项目 2 项,精品科普期刊项目 5 项,共 199 项。《中国中西医结合杂志》进入 TOP50 项目。