

• 循证医学与中医 •

临床试验结果的统计分析

万霞 邢建民 刘建平

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

[摘要] 统计分析方法在临床试验中起着十分重要的作用, 从临床试验的设计、监督、中期分析、终末分析到最后的报告, 统计分析方法贯穿着临床试验的始终。主要介绍了临床试验统计分析报告包含的内容、统计分析集的类型及应用、临床试验中常见的设计方法及常用的统计分析方法, 并举例说明统计分析中常见的问题。

[关键词] 统计描述; 统计推断; 统计方法

临床试验结果数据的分析要根据研究目的、研究设计方案和观测资料的类型来选择适当的统计模型, 并根据统计学原理决定应当采用参数统计还是非参数统计分析^[1]。临床试验结果的统计分析有 3 个目的: 质量控制、安全监测和试验结束后的统计分析报告^[2]。本文以试验结束后的统计分析报告中所涉及的统计分析要求, 来说明临床试验中常用的统计分析方法。

1 临床统计分析报告包含的内容

统计分析报告的内容包括: (1) 对整个临床试验中资料收集和整理过程的描述: 包括观察病例是否符合入选、排除标准; 各试验中心是否按照试验设计方案规定的观察病例数完整地收集到相应的数据; 试验过程中有无增加新的观察指标; 对脱落病例的处理方法和理由; 盲法试验应对盲态审核予以说明; 在资料整理过程中是否对指标进行数据变换; 如何定义离群值等内容。(2) 统计分析方法的选择及其理由: 包括统计分析所采用的方法, 例如: 疗效指标为等级变量, 两组比较采用行平均得分法; 对于等级资料, 扣除多中心效应采用 CMH χ^2 检验。(3) 各组病例入选及完成情况: 主要包括入选病例、完成情况、剔除和脱落的例数, 并列表报告脱落、剔除病例的原因, 各种统计分析数据集的例数。(4) 统计描述、参数估计、统计检验: 各组病例的主要指标、次要指标和全局评价指标的统计描述、参数估计、统计假设检验及优效性、非劣性或等效性的统计检验。(5) 各组病例安全性评价: 包括不良反应发生率及不良事件的具体描述; 实验室检查结果试验前后的变化情况, 以及发生异常改变时与试验药物的关系。

2 统计分析集的类型及应用

2.1 统计分析集的类型^[3]

常见的统计分析数据集有意向性治疗 (Intention To Treat, ITT) 分析集、全分析集 (Full analysis set,

FAS)、符合方案集 (Per-protocol population, PP) 和安全集 (Safety set, SS)。(1) ITT 分析集: 是指对所有随机化并分配到各个处理组的病例进行评价和分析, 而不论其是否依从计划完成治疗^[4]。这种方法在实际工作中很困难。例如, 在随机化后发现患者不符合主要的人组标准, 患者一次也没有用过试验药物或是在随机化后再没有任何资料等, 对这些病例进行分析显然不合理。所以 ITT 分析实际上可以说仅仅是一个原则。(2) FAS: 尽可能完整和尽可能按 ITT 原则包括所有随机化了的病例, 尽可能少和尽可能公平地排除病例^[5]。将其中未能观察到全部治疗过程的病例中的疗效指标, 用最后一次观察到的结果往后推的 CFLO (carry forward of the last observation) 原则。显然, FAS 的结果往往会低估疗效。(3) PP: 指 FAS 中更加符合方案的病例。什么情况把病例排除出符合方案集, 一般是在统计分析前由研究者根据试验方案确定, 并以书面形式加以记录。由于未能按方案完成治疗者排除在分析之外, 就使分析所得疗效较实际偏高。(4) SS: 指曾经应用其所在组药物至少 1 次者, 用于安全性分析。为了能全面反映药物的安全性, 在进行安全性分析时, 只要患者用过至少 1 次药物, 就应当研究其不良事件或不良反应, 不管此患者是否包括在 PP 中。

2.2 统计分析集的应用

在优效性试验 (验证试验药优于对照药的试验)、等效性试验 (验证两种药物疗效相等的试验) 或非优效性试验 (验证试验药是否不劣于对照药的试验) 中, FAS 和 PP 的作用不同。在优效性试验中为了避免由于 PP 对疗效的偏大估计, 一般以 FAS 作为主要的分析集, 因为它比较保守。而在等效或非优效性试验中用 FAS 一般并不保守。

3 临床试验中常见的设计方法

在临床试验中, 把试验对象按“完全随机设计”、“配对设计”或“随机区组设计”, 不同的设计方法不仅选择试验对象的方法不同, 在处理其结果时使用的统

基金项目: 国家重点基础研究发展计划资助项目 (2006CB504602)

通讯作者: 万霞, wanxiasnake@163.com, (010) 64286757

计方法也不同^[6]。(1)完全随机设计:是将受试对象随机分配到各个处理组或对照组中,或分别从不同总体中随机抽样进行研究。它可以是两样本比较,也可以是多样本比较;各样本含量可以相等,也可以不等,但也不宜差别太大。(2)配对设计:是将受试对象按一定条件配成对子,称为配对样本,将它们随机分到两个组中,分别给予不同的处理。配对条件是指有可能影响实验效应的非主要因素。在人群试验中,常将性别相同,年龄、生活劳动条件相近的两个人配成对子。这样可增加各处理组之间的可比性。(3)随机区组设计:也称配伍组设计,是配对设计的扩展。配对设计的每一“对子”分别随机分到两个处理组中,而配伍组设计中的每个“配伍组”包含多个受试对象,要将它们分别随机分到各处理组。

4 常用的统计分析方法

4.1 统计描述

很多人误以为统计描述的方法只有均数和标准差。其实,不同的资料类型有不同的描述方法,不同的数据分布状态也有不同的指标。(1)计量资料:正态分布的资料采用均数和标准差,非正态分布的资料则可以采用中位数和四分位间距进行描述。(2)计数资料:可以采用率、构成比或相对比进行描述。

4.2 统计推断

假设检验(hypothesis test)和参数估计(parameter estimation)是统计推断的重要组成部分。假设检验是对总体的参数和分布先作无效假设,然后根据样本对总体提供的信息,推断是否拒绝或不拒绝无效假设。根据分布已知与否,假设检验有参数法和非参数法。

常用的统计推断方法可以分为计数资料常用的统计分析方法(见图 1)和计量资料常用的统计分析方法(见图 2)。

5 统计分析中常见的问题

5.1 t 检验的误用

在临床资料分析中,由于对配对设计和完全随机设计两种设计类型区别不开,常误用配对 t 检验和完全随机设计 t 检验。

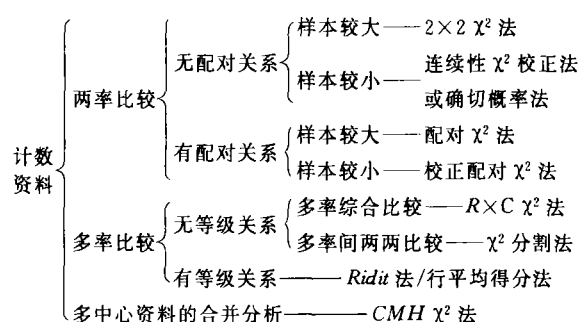


图 1 计数资料常用的统计分析方法

[例 1]原文题目:西药配合黄芪注射液治疗早期糖尿病肾病 30 例观察

治疗方法:60 例随机分为治疗组和对照组各 30 例。两组均采用糖尿病饮食控制,控制血糖用拜糖苹 50~100mg/d,或皮下注射胰岛素。治疗组另用黄芪注射液 30ml 静脉滴注,每日 1 次,连用 1 个月。两组的其他治疗如降压利尿等相同,两组均连用 1 个月为 1 个疗程。

观察方法:分别于治疗前和治疗后 1 个月检测尿微量白蛋白(Ualb)、血糖、血甘油三酯(TG)、胆固醇(Tch)。所有数据均采用 $\bar{x} \pm s$ 表示,配对 t 检验,以 $P < 0.05$ 为有统计学意义。

治疗结果:两组患者治疗前后的血糖、Ualb、TG、Tch 等参数的比较,见表 1。

表 1 两组治疗前后血糖、Ualb、血 TG、Tch 等参数比较($\bar{x} \pm s$)

组别	例数	血糖 ($\mu\text{mol/L}$)	Ualb ($\mu\text{g/ml}$)	Tch (mmol/L)	TG (mmol/L)
治疗组 30	治疗前	9.1 $\pm$ 1.2	6.7 $\pm$ 1.0	8.1 $\pm$ 0.8	2.8 $\pm$ 0.9
	治疗后	5.8 $\pm$ 1.4*	2.1 $\pm$ 0.8** Δ	6.6 $\pm$ 0.6	1.8 $\pm$ 0.7
对照组 30	治疗前	8.9 $\pm$ 1.5	6.9 $\pm$ 1.3	7.5 $\pm$ 0.5	2.5 $\pm$ 0.7
	治疗后	5.7 $\pm$ 1.7*	5.3 $\pm$ 1.1	6.1 $\pm$ 0.6	1.4 $\pm$ 0.6

与治疗前比较,* $P < 0.05$,** $P < 0.01$;与对照组比较, $\Delta P < 0.01$

对差错的辨析与释疑:原文统计方法是“所有数据均采用($\bar{x} \pm s$)表示,配对 t 检验,以 $P < 0.05$ 为有统计学意义”。本文并未交待按照某些因素配成对子,显然是一个完全随机设计,治疗组与对照组的比较采用配对 t 检验,是不正确的。但是对于治疗后与治疗前的比较,属于自身配对设计,应采用配对 t 检验的方法。

[例 2]原文题目:补阳还五汤加减治疗 2 型糖尿病早期肾

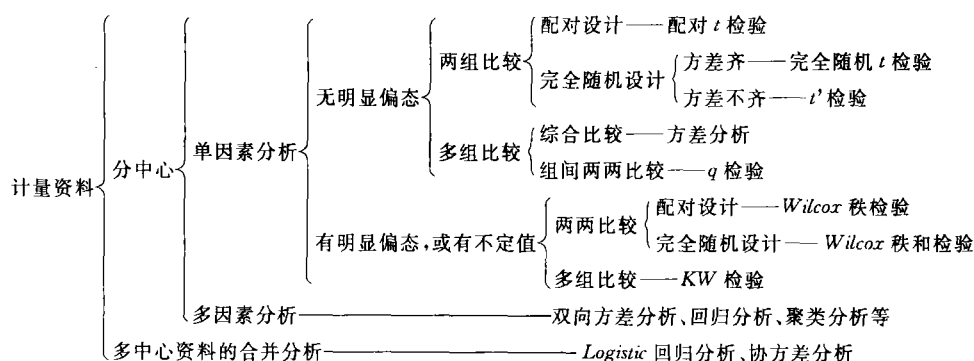


图 2 计量资料常用的统计分析方法

病的临床观察

治疗方法:将 85 例糖尿病肾病患者随机分为治疗组 43 例,对照组 42 例。治疗组予中药补阳还五汤,对照组不用任何改善肾功能的药物。两组均给予皮下注射胰岛素控制血糖。

观察项目:治疗前测空腹静脉血糖糖化血红蛋白(HbA1c)及 24 小时尿蛋白定量,治疗后每周复查上述指标 1 次。

统计学方法:所有数据均以均数±标准差($\bar{x} \pm s$)表示,各项计量资料用 t 检验。

结果:治疗组 43 例,失访 5 例,完成 38 例;对照组 42 例,失访 7 例,完成 35 例,两组治疗前后血糖、HbA1c 及 24 小时尿蛋白定量结果见表 2。两组治疗后比较 $P < 0.01$ 。

表 2 两组治疗前后血糖、HbA1c 及尿蛋白定量结果

组别	例数	空腹静脉血糖 (mmol/L)	HbA1c(%)	24 小时尿蛋白定量
治疗组				
治疗前	38	7.8±2.3	5.6±1.9	0.22±0.07
治疗后	38	7.2±1.7*	5.4±1.8△	0.13±0.02#
对照组				
治疗前	35	7.9±2.2	5.7±1.9	0.23±0.06
治疗后	35	7.5±2.1	5.5±1.7	0.24±0.05

治疗组治疗前后相比, * $P > 0.05$, △ $P > 0.05$, # $P < 0.01$

对差错的辨析与释疑:原文统计方法是“所有数据均以均数±标准差($\bar{x} \pm s$)表示,各项计量资料用 t 检验”。本文在表 2 中治疗组治疗前后相比,属于自身配对设计,应采用配对 t 检验的方法。而原作者只交待采用 t 检验,因此读者无法判断是采用配对 t 检验还是完全随机设计 t 检验。

5.2 等级资料分析不恰当

在临床资料中,对于等级资料的统计处理存在一个通病,许多作者都将等级资料转换成二分类变量进行处理,这样做会损失信息。具体实例如下:

[例 3]原文题目:中西医结合治疗糖尿病肾病的临床观察

治疗方法:160 例均为我院住院患者,随机分为两组。治疗组和对照组各 80 例。两组均予降血糖(诺合龙、糖适平及诺合灵)、调血脂、降血压(ACEI、血管紧张素 II 受体拮抗剂及钙离子拮抗剂)等常规治疗。治疗组中加用中药治疗。

统计学方法:计量资料采用 t 检验,计数资料采用 χ^2 检验。

治疗结果:疗效见表 3。

表 3 两组临床疗效比较(例)

组别	例数	显效	有效	无效	总有效率
对照组	80	31	43	6	92.4%*
治疗组	80	16	38	26	67.5%*

与对照组比较, * $P < 0.01$

对差错的辨析与释疑:原文统计方法中交待了“计数资料采用 χ^2 检验”。本文临床疗效指标为等级资料,在文中进行统计处理时,将原始资料分为了两类“有效”(即“显效”+“有效”)和“无效”,因此,采用了 χ^2 检

验。这样的处理方法,使得对照组与治疗组中的“显效”和“有效”两个级别的差异被掩盖了。因此,正确的处理方法应当是采用“Ridit 分析”或者“行平均得分法”,分析两组间 3 个级别(显效、有效和无效)之间是否有显著性差异。若有差异,再考虑做 χ^2 分割,进一步比较各个等级间是否有差异。

6 结语

统计分析方法在临床试验中起着十分重要的作用,从临床试验的设计、监督、中期分析、终末分析以及最后的报告,统计分析方法贯穿着临床试验的始终^[7]。但是,统计方法涉及面很广,各有其适用范围,不可误用,它决定了临床试验结论的真实性。因此,临床研究人员只有很好地掌握了统计学知识,才能合理运用。为了保证结果的正确性,统计分析还需由两个程序员独立地编制统计分析程序,并且获得同样的计算结果。只有这样,临床试验的结果才能正确、可靠,结论才可信。

参考文献

- [1]苏炳华主编. 新药临床试验统计分析新进展. 上海:上海科学技术文献出版社, 2000: 40.
- [2]Meinert CL, Tonascia S. Clinical Trials design, conduct, and analysis. Oxford University Press. 1986: 179.
- [3]金丕焕. 医学科学研究中的数据管理和分析集. 中华预防医学杂志, 2002, 36(1): 68~70.
- [4]Intention-to-treat analysis (ITT). Bandolier EB/OL. [2007-10-21] <http://www.jr2.ox.ac.uk/bandolier/booth/glossary/ITT.html>
- [5]ICH Harmonised Tripartite Guideline Statistical Principles for Clinical Trials E9. (Current Step 4 version dated 5 February 1998). EB/OL. [2007-10-21] <http://www.ich.org/LOB/media/MEDIA485.pdf>
- [6]马斌荣. 医学统计学. 4 版. 北京:人民卫生出版社, 2004: 6.
- [7]Statistics in clinical trials. Children's cancer and leukaemia group. EB/OL. [2007-10-21] http://www.cclg.org.uk/public/clinical_trials/13stats.html

(收稿日期:2007-11-15)

编者按:循证医学的理念和方法引入中医药的临床评价研究,必将促进中医药临床实践从经验医学向以证据为基础的临床实践发展,促进中医现代化进程。本刊从 2007 年第 1 期开始,特设了“循证医学与中医”专栏,详细介绍了如何采用循证医学的方法进行中医疗效评价,阐述了在临床试验中如何设置对照、设立盲法、如何进行随机分组、临床试验的样本量估算、如何选取结局指标并进行测量,如何进行病例报告表的设计。特别针对针灸试验的特殊性,介绍了针灸临床试验报告的国际标准——STRICTA 指南。为促进正确地分析和报告中医临床研究结果,还介绍了临床资料的统计分析方法及常见错误分析。希望读者对本栏目系列文章提出意见和建议,以便今后给读者带来更加有益的循证医学方法和相关信息。