

【临床研究】

完全随机设计和配对设计临床试验研究的统计分析方法及比较

王瑞平¹, 陈 曦², 陈 洁², 李 欣², 强 燕³, 付晓伶¹, 李 斌^{2,4*}

(1.上海中医药大学附属岳阳中西医结合医院临床研究中心办公室, 上海 200437; 2.上海中医药大学附属岳阳中西医结合医院皮肤科, 上海 200437; 3.上海交通大学医学院附属松江医院皮肤科, 上海 201620; 4.上海市中医药研究院皮肤病研究所, 上海 201203)

摘要: **目的** 分析临床试验中完全随机设计和配对设计在数据整理和统计分析过程的差异, 提高临床研究人员对两种设计方案的认知和研究分析能力。**方法** 基于银屑病“凉血潜阳法”临床研究选取100例对象, 分别按照“完全随机设计”和“配对设计”的方案进行数据整理, 建立模拟数据库, 对不同设计的数据库架构、统计分析指标选择和统计分析等内容进行比较分析。**结果** 完全随机设计通过“组别”变量区分试验组和对照组, 配对设计通过在变量名上添加(1, 0)区分试验组和对照组, 变量个数翻倍。完全随机设计和配对设计在统计分析指标选择、统计描述指标、单因素分析和重复测量数据统计分析、数据展示方式和图形选择等方面均存在区别。**结论** 临床试验研究中, 应综合考虑研究设计类型和研究结果的数据分类, 选择合适的统计分析模型和指标, 改善临床试验中统计分析结果的准确性。

关键词: 完全随机设计; 配对设计; 临床试验; 比较分析

中图分类号: R181.2 **文献标志码:** A **文章编号:** 1672-9188(2019)11-0783-07

DOI: 10.13683/j.wph.2019.11.007

Comparison of statistical analysis between clinical trials with paired design and simple randomized design

WANG Rui-ping¹, CHEN Xi², CHEN Jie², LI Xin², QIANG Yan³, FU Xiao-ling¹, LI Bin^{2,3*}

(1. Office of Clinical Research Center, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai 200437, China; 2. Department of Dermatology, Yueyang Hospital of Integrated Traditional Chinese and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai 200437, China; 3. Department of Dermatology, Songjiang Hospital, Medical College, Shanghai Jiaotong University, Shanghai 201620, China; 4. Research Institute of Dermatology, Shanghai Academy of Traditional Chinese Medicine, Shanghai 201203, China)

Abstract: **Objective** To compare the difference of clinical trials with paired design and simple randomized design in the aspect of data arrangement and data analysis, and improve the recognition of clinical trial design and data analysis ability among clinical trial practitioners. **Methods** Based on a psoriasis clinical trial titled 'Cooling Blood and Latent Yang Method', 100 patients were extracted and rearranged into two simulation databases in accordance with the study plan of 'simple randomized design' and 'paired design', then we compare the differences of database structure, statistical index selection and statistical analysis between clinical trials with the paired design and simple randomized design. **Results** In RCT with simple randomized design, a variable named 'group' is used to differentiate the treatment group and control group, whereas in RCT with paired design, the number of variables are doubled with each pair added a '1' or '0' to differentiate the treatment group and control group. Compared with RCT with simple randomized design, the statistical index selection, descriptive analysis, single factor analysis, repeated measurement analysis, data visualization method in RCT with paired design is different.

收稿日期: 2019-06-27

作者简介: 王瑞平, 博士, 副主任医师, 研究方向: 临床流行病学设计和统计方法学。

通讯作者: 李斌, 博士, 主任医师, 教授, 研究方向: 中医药防治皮肤病。

基金项目: 国家重点研发计划“中医药现代化研究”重点专项(2018YFC1705301); 上海市松江区科委科技攻关项目(18SJKGG58)

Conclusion In clinical trials, it is critical to contemplate the study design and data type comprehensively before data analysis, and then select the proper statistical analysis model and statistical indexes accordingly could improve the correctness and accuracy of data analysis in clinical trials.

Key words: simple randomized design; paired design; clinical trial; comparison analysis

临床试验是指任何在人体进行的药物系统性研究,以证实或揭示试验药物的作用、不良反应和/或(或)试验药物的吸收、分布、代谢及排泄,目的是确定试验药物的疗效和安全性^[1-2]。临床试验的核心是随机分组和设立对照组,因此随机对照试验(randomized control trial, RCT)是一种经典的临床试验类型^[3]。临床试验方案的设计是临床试验顺利开展的重要保障,其不仅关系到试验对象的选择方法,同时也决定后续统计分析方法的选择,是一个不可忽视的重要环节^[4-5]。“完全随机设计”和“配对设计”是RCT研究中最常见的两种类型,近年来在临床试验中得到广泛应用。鉴于数理统计分析的专业性,部分临床研究人员对完全随机设计和配对设计RCT统计分析方法的理解存在混淆,导致分析结果的错误^[6]。本文拟通过基于同一批银屑病“凉血潜阳法”RCT研究数据,分别按完全随机设计和配对设计方案模拟数据整理和统计分析过程,阐释比较两种不同设计方案的统计分析结果和差异,以期改善临床研究人员对两种设计方案临床研究数据分析过程的认知,提高临床研究分析能力和水平。

1 资料与方法

1.1 资料来源

本研究资料选自促进市级医院临床技能和临床创新能力三年行动计划“银屑病‘凉血潜阳法’临床方案优化及再评价研究”。从该研究中分别随机选取50例试验组和50例对照组为研究对象,摘录变量主要为年龄、性别、银屑病皮损面积和严重程度指数(psoriasis area and severity index, PASI)评分及治疗效果等指标。

1.2 数据模拟

基于上述银屑病“凉血潜阳法”临床研究中选

取的100例对象,分别按照“完全随机设计”和“配对设计”的方案设计原理,对摘选数据进行整理。完全随机设计是将样本中全部受试对象随机分配到各个处理组,分别接受不同的处理,然后对效应进行比较。将上述50例试验组和50例对照组对象的资料进行纵向合并,增加一个新的变量名“组别”Group(1=试验组,0=对照组),建立完全随机设计的模拟数据库A(样本量 $n=100$);配对设计是将受试对象按某些特征或条件配成对子,再将每对中两个受试对象随机分配到两个不同处理组中接受不同的处理,然后比较疗效和安全性。将上述50例试验组和50例对照组对象的资料进行横向合并,以银屑病患者PASI评分及年龄进行个体匹配,建立配对设计的模拟数据库B(样本量 $n=50$)。

1.3 统计分析

应用SPSS16.0软件,示例分析完全随机设计和配对设计资料。主要包括一般人口学特征描述,应用 t 检验(配对 t 检验)或Wilcoxon秩和检验分析定量变量资料差异;应用Pearson卡方检验或McNemar卡方检验分析定性变量资料差异;应用广义线性模型广义线性模型(general linear model, GLM)分析重复测量数据,检验水准 $\alpha=0.05$,以 $P<0.05$ 为差异有统计学意义。

2 结果

2.1 不同研究设计的数据库架构特点

临床研究中,完全随机设计通过“组别”变量区分试验组和对照组。如表1所示,完全随机设计的银屑病RCT模拟数据库中共包括100例银屑病患者,组别、性别、年龄、体表面积(body surface area, BSA)、PASI0-PASI8和疗效均为独立变量,通过控制“组别”变量可以分析不同组别对象各指标

之间的差异。配对设计中, 受试对象先按照一定条件配成对子, 再将每对中两个受试对象随机分配到两个不同处理组中接受处理。如表 2 所示, 配对设计的银屑病 RCT 模拟数据库中共包括 50 例银屑病患者, 组别、性别、年龄、BSA、PASI0-PASI8 和疗效均为相关变量, 通过在变量名上添加 (1, 0) 来区分试验组和对照组, 统计分析中通过计算对应相关变量的差值 (定量变量) 或不一致的对子数 (定性变量) 来比较不同组别的差异有无统计学意义, 详见表 1 和表 2。

表 1 基于完全随机设计的银屑病
RCT 研究模拟数据库架构

Table 1 Database structure of RCT study for psoriasis
based on the simple randomized design

ID	组别	性别	年龄	BSA	PASI0	...	PASI8	疗效
1	1	1	28	5	1.10	...	1.00	4
2	1	1	67	1	1.20	...	1.00	4
...
50	1	1	69	1	2.90	...	3.00	4
51	0	2	48	1	0.30	...	0	4
52	0	2	63	5	1.20	...	0.40	3
...
99	0	2	26	1	3.60	...	2.80	4
100	0	2	38	10	3.60	...	2.10	3

2.2 不同研究设计的分析指标选择

RCT 研究中, 完全随机设计和配对设计的统计分析指标选择存在一定的区别。例如, 在完全随机设计 RCT 研究中, 定量资料如果符合正态分布, 两样本之间的 t 检验公式为 $t = (\bar{x}_1 - \bar{x}_2) / \sqrt{s_c^2(1/n_1 + 1/n_2)}$, 其中 $S_c^2 = [(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2] / (n_1 + n_2 - 2)$, 根据两样本的方差齐性检验结果, 选择 t 检验或者 t' 检验。而配对设计 RCT 研究中, 定量资料如果符合正态分布, 两样本之间的 t 检验公式为 $t = (\bar{d} - 0) / (s_d / \sqrt{n})$, 其检验的是两样本资料某指标差值 (\bar{d}) 总体均数是否为 0, 配对 t 检验的原理与完全随机设计 t 检验不同, 应根据不同的研究设计选择恰当的统计分析指标开展统计工作。图 1 汇总了完全随机设计和配对设计 RCT 研究的统计分析方法, 研究者可根据不同的研究设计和资料类型选择合适的方法。

2.3 不同研究设计的统计学描述

在 RCT 研究主要效应指标分析前, 需要对研究对象的人口学资料、实验室检测结果资料和试验效应指标进行统计学描述分析。定量变量资料选择 $(\bar{x} \pm s)$ (正态或近似正态分布) 或中位数和四分位间距 (偏态分布) 描述。如表 3 所示, 完全随机设计描述全体对象的变量特征, 配对设计分别描述试

表 2 基于配对设计的银屑病 RCT 临床研究模拟数据库架构

Table 2 Database structure of RCT study for psoriasis based on the paired design

对子	试验组							对照组						
	性别1	年龄1	BSA1	PASI10	...	PASI18	疗效1	性别0	年龄0	BSA0	PASI00	...	PASI08	疗效0
1	1	28	5	1.10	...	1.00	4	2	48	1	0.30	...	0	4
2	1	67	1	1.20	...	1.00	4	2	63	5	1.20	...	0.40	3
3	1	31	45	1.80	...	0.60	3	1	55	5	1.60	...	0.60	3
4	2	55	2	1.80	...	4.20	4	2	40	8	1.90	...	2.10	4
...
46	1	30	4	2.40	...	1.40	3	2	61	12	3.00	...	1.00	3
47	2	68	10	2.70	...	2.00	4	1	50	50	3.45	...	1.40	3
48	1	57	6	2.80	...	1.60	3	1	28	3	3.50	...	3.00	4
49	1	27	1	2.90	...	2.80	4	2	26	1	3.60	...	2.80	4
50	1	69	1	2.90	...	3.00	4	2	38	10	3.60	...	2.10	3

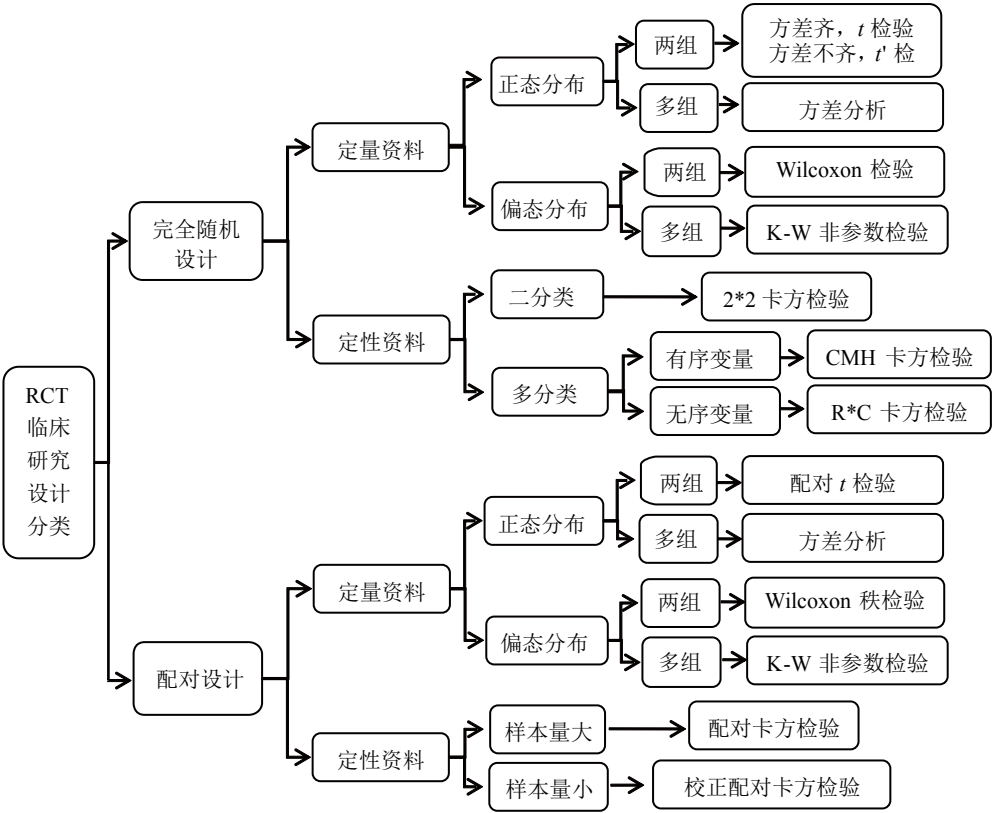


图 1 RCT 研究不同设计方案的统计分析思路图

Figure 1 Flowchart of statistical analysis for RCT based on different study design

表 3 不同研究设计临床试验研究中常用的定量变量统计学描述指标值

Table 3 Statistical indexes of descriptive analysis for quantitative variable in RCT with different study design

设计 分类	变量名	样本量	统计量						K-S	P
			平均值	标准差	最小值	最大值	中位数	四分位间距		
完全随 机设计	年龄/岁	100	45.44	13.32	20.00	69.00	46.00	23.00	0.09	0.05
	BSA	100	35.10	28.75	1.00	90.00	32.50	55.00	0.17	<0.05
	PASI0	100	11.16	9.44	0.30	38.40	7.80	13.30	0.17	<0.05
	PASI8	100	4.88	4.66	0.00	23.00	3.40	5.00	0.16	<0.05
配对 设计	年龄0/岁	50	46.54	14.41	20.00	69.00	47.50	26.00	0.11	0.20
	BSA0	50	34.80	31.06	1.00	90.00	30.00	58.00	0.18	<0.05
	PASI00	50	11.43	9.30	1.10	38.30	9.30	14.60	0.14	<0.05
	PASI08	50	5.74	5.03	1.00	23.00	4.45	6.00	0.15	<0.05
	年龄1/岁	50	44.34	12.18	22.00	67.00	44.00	22.00	0.12	0.08
	BSA1	50	35.40	26.56	1.00	90.00	40.00	50.00	0.17	<0.05
	PASI10	50	10.88	9.67	0.30	38.40	6.88	9.20	0.24	<0.05
	PASI18	50	4.02	4.14	0.00	20.00	2.80	3.00	0.20	<0.05

验组和对照组的变量特征。根据正态性检验结果, 银屑患者的年龄用 $(\bar{\mu} \pm \sigma)$ 描述, BSA、PASI0 和 PASI8 用中位数和四分位间距描述。定性变量应用构成比 (percentage, %) 或率 (rate, %) 描述, 如表 4 所示, 选择构成比描述研究对象的性别和治疗效果。

表 4 不同研究设计临床试验研究中常用的定性变量统计学描述指标值

Table 4 Statistical indexes of descriptive analysis for qualitative variable in RCT with different study design

变量名	完全随机设计		配对设计			
			试验组		对照组	
	人数/例	构成比/%	人数/例	构成比/%	人数/例	构成比/%
性别						
男	66	66.00	26	52.00	40	80.00
女	34	34.00	24	48.00	10	20.00
治疗疗效						
显效	26	26.00	7	14.00	19	38.00
有效	51	51.00	29	58.00	22	44.00
无效	23	23.00	14	28.00	9	18.00

2.4 不同研究设计的单因素分析和结果展示

完全随机设计 RCT 研究指标的单因素比较中, 定量变量根据正态性检验结果, 采用 t 检验或 Wilcoxon 秩和检验比较不同组的差异。如表 5 所示, 分别描述试验组和对照组银屑患者的统计指标值 $[(\bar{\mu} \pm \sigma)$ 或平均秩次], 并选择合适的统计量进行检

验结果的表达。配对设计 RCT 研究指标的单因素比较中, 根据正态性检验结果, 采用配对 t 检验或两相关样本秩和检验比较试验组和对照组不同指标差值与 0 之间的差异。如表 5 所示, 符合正态分布用指标差值的 $(\bar{\mu} \pm \sigma)$ 表示, 不符合正态分布用正秩次和负秩次表示, 并选择合适的统计量进行检验结果的表达。RCT 研究定性变量分析中, 完全随机设计采用构成比 (%) 展示不同组对象指标的构成情况, 应用 Pearson 卡方检验进行统计分析, 见表 6; 配对设计中分别统计试验组和对照组在不同指标之间的对子数和构成情况 (%), 应用 McNemar 配对卡方检验进行统计分析, 见表 7。

表 6 完全随机设计临床试验中定性变量的单因素分析
Table 6 Univariate analysis for qualitative variable in RCT with simple randomized design

治疗结果	试验组		对照组		χ^2_{Pearson}	P
	人数/例	构成比/%	人数/例	构成比/%		
显效	19	38.00	7	14.00	7.59	<0.05
有效	22	44.00	29	58.00		
无效	9	18.00	14	28.00		

2.5 不同研究设计的重复测量数据分析和结果展示

RCT 研究中, 为动态观察效应指标的变化情况, 通常会对效应指标进行多个时间点的检测。例如在银屑病“凉血潜阳法”临床研究中, 对每一例银屑

表 5 不同研究设计的临床试验研究中定量变量的单因素分析

Table 5 Univariate analysis for quantitative variable in RCT with different study design

变量	组别	完全随机设计				变量	配对设计(差值)				
		$\bar{\mu} \pm \sigma$	平均秩次	t/Z	P		$\bar{\mu} \pm \sigma$	正秩次	负秩次	$t/K-S$	P
年龄/岁	试验组	44.34 ± 12.18	—	0.83	0.41	年龄/岁	2.20 ± 2.83	—	—	0.78	0.44
	对照组	46.54 ± 14.41	—			BSA	—	22.54	23.52	-0.27	0.79
BSA	试验组	—	51.82	0.46	0.65	PASI0	—	20.46	29.02	-1.41	0.16
	对照组	—	49.18			PASI8	—	24.05	23.35	3.28	<0.05
PASI0	试验组	—	49.84	-0.23	0.82						
	对照组	—	51.16								
PASI8	试验组	—	44.35	-2.12	<0.05						
	对照组	—	56.65								

表 7 配对设计临床试验中定性变量的单因素分析[n(%)]

Table 7 Univariate analysis for qualitative variable in RCT with paired design [n(%)]

对子数	对照组			χ^2 McNemar	P
	显效	有效	无效		
显效	3 (6.00)	3 (6.00)	1 (2.00)	7.60	0.06
试验组 有效	12 (24.00)	13 (26.00)	4 (8.00)		
无效	4 (8.00)	6 (12.00)	4 (8.00)		

病患者的 PASI 评分进行了 0 周、2 周、4 周、6 周和 8 周共 5 次评估。把重复测量数据作为独立组别进行方差分析是不准确的, 应选择 GLM 重复测量数据进行方差分析。如表 8 所示, 完全随机设计需分析实验组别、处理时间、组别和时间交互作用的效应; 而配对设计由于将试验组和对照组进行了配对, 只需分析处理时间的效应。GLM 分析中, 应采取 Mauchly 球形检验分析重复测量数据整体上是否符合 Huynh-Feldt 条件, 如果拒绝了 H_0 球型假设, 需要

进行 Greenhouse-Geisser (G-G) 或 Huynh-Feldt (H-F) 调整 P 值。对于重复测量资料分析结果的图形展示选择, 应考虑分析指标是否符合正态分布。如图 2 所示, 如果不同周测量的 PASI 评分符合正态分布, 选择图 2-Part B 部分所使用的直条图与误差线的组合; 如果不符合正态分布, 选择图 2-Part A 部分所使用的箱式图。

3 讨论

生物统计是运用数理统计的原理, 结合医学和生物学实际, 研究数字资料的收集、整理、分析和推断的一门学科, 是进行临床研究所必须的重要手段^[7]。临床试验结果数据的分析要根据研究目的, 研究设计方案和观察资料的类型选择合适的统计分析模型, 并根据统计学原理决定采用参数统计还是非参数统计方法^[8]。本文从临床研究中最常见的两种设计类型出发, 通过基于同一批银屑病“凉血潜

表 8 不同研究设计临床试验中定量变量的重复测量数据方差分析

Table 8 ANOVA for repeated measured qualitative variable in RCT with different study design

研究设计	变异来源	自由度	离均差平方和	均方	F	P	调整P值	
							G-G	H-F
完全随机设计	实验组别	1	153.26	153.26	0.79	0.38		
	组间误差	98	18957.88	193.45				
	处理时间	4	2603.24	650.81	61.15	<0.05	<0.05	<0.05
	组别与时间交互项	4	18.16	4.54	0.43	0.79	0.59	0.59
	组内误差	392	4171.77	10.64				
配对设计	处理时间	4	36.31	9.08	1.54	0.19	0.22	0.21
	误差	196	1155.30	5.89				

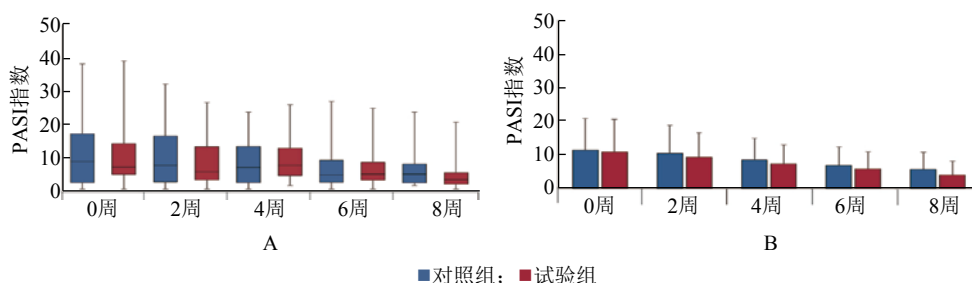


图 2 RCT研究分析指标的绘图选择示例

Figure 2 Charts for RCT based on different analysis indexes

阳法” RCT 数据, 分别按完全随机设计和配对设计方案模拟数据整理和统计分析过程, 阐释比较两种设计方案的统计分析结果和差异, 对于提高临床研究人员对不同设计类型的数据整理和统计分析有所帮助。

临床研究设计中, 完全随机设计是将样本中全部受试对象随机分配到各处理组, 分别接受不同的处理, 然后对其效应比较。配对设计是将受试对象按照某些特征或条件配成对子, 再将每对中两例受试对象随机分配到两个不同处理组中, 然后对其效应比较。完全随机设计和配对设计各有优势, 完全随机设计便于实施, 而配对设计的统计学效率更高。需要注意的是, 配对设计中配对因素一般选择那些影响试验效应的非处理因素, 配对因素一般选择 1~2 个; 同时由于配对因素在试验组和对照组进行了匹配, 后续的统计分析中将不能分析配对因素对试验结果的影响, 临床研究人员在选择设计方案时应进行认真考虑。

在数据库架构方面, 不同统计分析软件对变量的设置要求有所区别。对于完全随机设计, 常用的统计分析软件如 SPSS、SAS、Stata、Epi-info 等对数据库架构要求相同, 均通过变量“组别”来区分不同的处理组, 后续的分析策略和分析方法选择也一致。但对于配对设计, SPSS、SAS 和 Stata 分析软件对数据库架构要求相同, 通过在所有变量名上添加 (1, 0) 来区分试验组和对照组, 变量名数量翻倍。统计分析中通过计算对应相关变量的差值 (定量变量) 或不一致的对子数 (定性变量) 来比较不同组别的差异有无统计学意义。对于配对设计, Epi-info 软件对数据库架构要求与 SPSS、SAS 和 Stata 分析软件不同, Epi-info 软件通过在完全随机设计数据库的基础上增加一个“对子”变量名实现对不同组别中个体之间的配对, 在分析过程中将“对子”变量进行控制, 实现配对设计的统计分析。

统计分析方法和指标的选择取决于试验结果的数据类型。临床试验数据可以分为两大类, 即定量

变量数据和定性变量数据^[9]。对于定量变量数据, 主要描述数据的集中趋势和离散趋势两个特征, 具体统计学指标的选择取决于定量变量数据的正态性检验结果。如果定量变量数据符合正态分布, 用均数和标准差描述, 采用 *t* 检验和方差分析进行组间效果比较。如果不符合正态分布, 用中位数和四分位间距描述, 采用 Wilcoxon 秩和检验进行组间效果比较。对于定性变量数据, 主要应用构成比或率描述, 采用 Pearson 卡方检验、CMH 卡方检验进行组间效果比较。同时, 对于不同的设计类型, 如本文中提及的完全随机设计和配对设计, 对定量变量数据和定性变量数据的分析方法也不相同。因此, 临床研究人员在开展数据统计分析之前, 应认真思考研究设计类型和数据类型, 选择合适的统计分析方法, 提高研究结果的可靠性。

临床研究中, 研究对象失访 (脱落) 会对研究结果造成一定的影响。因此, 在临床研究开展的各个阶段均应实施控制措施, 避免研究对象的失访, 将失访率控制在 15% 以内, 同时详细记录失访对象的特征和失访原因。数据分析过程中, 应比较失访对象和完成临床研究对象在一般人口学特征 (年龄、性别等)、试验组别及疾病的一般特征等方面是否存在差异。如果失访对象与完成临床研究对象一般人口学特征、疾病特征等方面无差异, 同时试验组和对照组的失访率无显著性差异, 那么可以认为失访对整体研究和结论影响不大; 否则应考虑失访对研究结果的影响, 慎重解读研究结论。

4 结论

临床试验研究中, 应综合考虑研究设计类型和研究结果的数据分类, 选择合适的统计分析模型和指标, 改善临床试验中统计分析结果的准确性。

参考文献:

- [1] Zhang X, Tian R, Yang Z, *et al.* Quality assessment of clinical trial registration with traditional Chinese medicine in WHO registries[J]. *BMJ Open*, 2019, 9(2): e025218-e025218.

(下转第 794 页)