

多中心临床随机对照试验的 Meta 分析

欧爱华 老膺荣 李晓庆 黄燕 何羿婷 黄培新

【摘要】 目的 探讨解决多中心临床随机对照试验中心效应差异的统计分析问题。方法 以两项多中心临床随机对照试验数据为例,运用协方差分析及 Meta 分析。结果 协方差分析中心间效应值差异均有统计学意义,行 Meta 分析,项目 1 异质性检验差异无统计学意义($P > 0.05$),采用固定效应模型分析合并效应值组间差异有统计学意义($P < 0.05$)。项目 2 异质性检验差异有统计学意义($P < 0.05$),采用随机效应模型分析合并效应值组间差异无统计学意义($P > 0.05$)。结论 多中心临床随机对照试验研究中,如果存在中心间以及中心与分组间交互效应差异有统计学意义时,可根据 Meta 分析异质性检验结果选择适合的模型进行合并效应值的组间比较,如果协方差与 Meta 分析结果不一致时,建议选择 Meta 分析的结果较为稳妥。

【关键词】 临床试验,多中心; Meta 分析; 中心效应

A Meta-analysis in multi-center random controlled clinical trials OU Ai-hua, LAO Ying-rong, LI Xiao-qing, HUANG Yan, HE Yi-ting, HUANG Pei-xin. Department of Clinical Epidemiology Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangzhou 510120, China

【Abstract】 Objective To study the center effect discrepancy in the multi-center clinical trials. **Methods** Two groups of data collected from the multi-center clinical trials were used. Data were processed by covariance analysis and Meta-analysis. **Results** In the covariance analysis, the discrepancy of the center effect values indicated statistical significance. Through Meta-analysis on fixed effect model, the discrepancy in one heterogeneity test showed no statistical significance ($P > 0.05$) while the inter-group discrepancy of the merged effect values drawn from analysis based on fixed effect model having statistical significance ($P < 0.05$). In the random effect model, the discrepancy in one heterogeneity test showed statistical significance ($P < 0.05$) while the inter-group discrepancy of the merged effect values drawn from analysis based on random effect model having no statistical significance ($P > 0.05$). **Conclusions** Studies on multi-center random controlled clinical trials, when statistical significance was found in the interaction discrepancy between the inter-center and the center-group relation, the merged effect values should be compared and analyzed by an appropriate statistic model based on the heterogeneous test results from the Meta-analysis. However, if the result from covariance analysis and the one from Meta-analysis did not agree to each other, the results drawn from the Meta-analysis were reliable.

【Key words】 Multi-center random controlled clinical trials; Meta-analysis; Center effect

在新药临床试验以及国家重大项目的临床试验研究中,由于研究项目涉及多个中心参与,尽管临床试验前制定了统一的临床试验方案,从研究目标、诊断标准、纳入标准、排除标准、疗效标准等均统一按临床试验方案实施,并在临床试验前对参与临床试验的研究人员进行了统一的培训,但仍然避免不了各中心疗效的不一致性(中心效应)等问题^[1],在分析中心效应的问题上,过去通常采用协方差分析解决,但有时也有一定的困难。近年来 Meta 分析作

为临床治疗性文献系统评价的有用分析工具之一^[2-5],它是对多个同类研究结果进行汇总及合并分析,从而达到增大样本含量,提高检验效能的目的,尤其在多个研究结果不一致或都无统计学意义时,用 Meta 分析可得到更加接近真实情况的分析结果。本文以实例,在协方差分析的基础上,同时采用 Meta 分析,并进行两种方法的比较和分析。

基本原理

1. 原理: Meta 分析的过程是进行参数的估计和假设检验,即得到效应合并值的点估计和区间估计,参数估计的方法与分析变异的来源和选择何种模型十分重要。因此,应对 Meta 分析效应合并时的

基金项目:国家“十五”科技攻关课题资助项目(2001BA701A17、2001BA701A12b)

作者单位:510120 广州,广东省中医院临床流行病学室

变异来源进行分析,其来源有两类:一类是抽样误差,即研究内变异(inner-study variation),它可能是由于各中心单独研究的样本含量不同,样本内的各观察单位可能存在差异,但当样本含量较大时,抽样误差相对较小。另一类是研究间变异(inter-study variation),它可能是由于研究对象以及偏倚的控制等诸多方面存在差异,造成研究的质量有高有低,如果研究间变异很小,可以认为研究间变异可能是由于抽样误差造成的,纳入 Meta 分析的各独立研究是来自同一总体,各独立研究间的效应只是效应合并值这一总体参数的估计值,此时可选择固定效应模型(fixed-effects model)。如果研究间变异较大,不能单用抽样误差来解释这种差异,应考虑各独立研究来自不同的、但又有联系的一些总体,每个研究有其相应的总体参数,Meta 分析的效应合并值是多个不同总体参数的加权平均,此时可选择随机效应模型(random-effects model)。

2. 模型:

(1)固定效应模型:假设各个独立研究的效应值为 Y_i ,相应的总体参数为 θ ,即 $E(Y_i) = \theta, s_i^2 = \text{var}(Y_i)$,当样本含量较大时,根据中心极限定理, Y_i 应该服从以 θ 为中心的近似正态分布。假定方差 s_i^2 已知,在固定效应模型中,有

$$Y_i \sim N(\theta, s_i^2) \tag{1}$$

$i = 1, 2, \dots, K$ 表示纳入的 K 个独立研究, θ 即为 Meta 分析的效应合并值。

(2)随机效应模型:假设 Y_i 为第 i 个独立研究的效应值, Y_i 来自均数为 θ_i ,方差为 s_i^2 的正态分布, θ_i 为第 i 个研究的“真正效应”, $\theta_1, \theta_2, \dots, \theta_k$ 相互独立,总体均数 θ 为效应合并值,总体方差 τ^2 为研究间变异,即随机效应,如假设 θ_i 来自均数为 θ ,方差为 τ^2 的正态分布,则在随机效应模型中,有

$$\begin{aligned} Y_i | \theta_i, s_i^2 &\sim N(\theta_i, s_i^2) \\ \theta_i | \theta, \tau^2 &\sim N(\theta, \tau^2) \end{aligned} \tag{2}$$

3. 参数估计:参数估计时,首先要作异质性检验,以便根据异质性检验的结果选择适合的模型。如果异质性检验不拒绝 H_0 ,可以认为研究间的差异无统计学意义,可选择固定效应模型;如果异质性检验差异有统计学意义,则拒绝 H_0 ,可采用随机效应模型。异质性检验

$$H_0: \theta = \theta_1 = \theta_2 = \dots = \theta_k$$

$$H_1: \theta_1, \theta_2, \dots, \theta_k \text{ 中至少有一个不全相同}$$

在 H_0 成立的条件下,对大样本研究,有统计量

$$Q_w = \sum_i^k w_i (Y_i - \theta)^2 \sim \chi_{k-1}^2 \tag{3}$$

由极大似然估计或加权最小二乘估计可得 θ 的估计值 $\hat{\theta}$,即

$$\hat{\theta} = \frac{\sum w_i Y_i}{\sum w_i} \quad w_i = \frac{1}{s_i^2}$$

如果 Q_w 不大于自由度为 $k-1$ 的 χ^2 分布的界值,则不拒绝 H_0 ,可以认为被纳入 Meta 分析的 K 个研究同质,即这些研究来自相同的总体,这时可选择固定效应模型;如果 Q_w 大于自由度为 $k-1$ 的 χ^2 分布的界值,则拒绝 H_0 ,可以认为被纳入 Meta 分析的 K 个研究不同质,即研究间存在异质性,这些研究可能来自 2 个或多个不同的总体,这时可选择随机效应模型。

4. 连续性数值变量的 Meta 分析方法:主要是以标准化的均数之差作为效应值,用 n_{1i}, n_{2i} 表示纳入 Meta 分析的第 i 个研究组和对照组的样本大小, $\bar{x}_{1i}, \bar{x}_{2i}$ 分别表示第 i 个研究组和对照组的均数, s_{1i}^2, s_{2i}^2 为第 i 个研究组和对照组的方差, s_{pi}^2 为两组合并方差,令 $y_i = \bar{x}_{1i} - \bar{x}_{2i}$,标准误为 s_i ,有

$$\begin{aligned} s_i^2 &= s_{pi}^2 (1/n_{1i} + 1/n_{2i}) \\ s_{pi}^2 &= \frac{(n_{1i} - 1)s_{1i}^2 + (n_{2i} - 1)s_{2i}^2}{n_{1i} + n_{2i} - 2} \end{aligned}$$

当纳入的研究结局指标量纲相同,此时可直接用两组的均数之差 y_i 作为效应,则效应的合并值为

$$y_{\text{合并}} = \frac{\sum w_i y_i}{\sum w_i} \quad w_i = \frac{1}{s_i^2}$$

异质性检验的统计量及效应合并值的 95% 可信区间(CI)分别为

$$\begin{aligned} Q &= \sum w_i (y_{\text{合并}} - y_i)^2 = \sum w_i y_i^2 - \frac{(\sum w_i y_i)^2}{\sum w_i} \\ y_{\text{合并}} &\pm 1.96 / \sqrt{\sum w_i} \end{aligned}$$

进行异质性检验计算效应合并值及效应合并值 95% CI 方法很多,有的计算较为复杂,一般可借助统计软件包实现其计算^[6]。以下是两项多中心临床试验研究实例分析。

实例分析

资料来源于广东省中医院国家“十五”科技攻关

项目,采用多中心临床单纯随机对照试验方法,试验组与对照组按 1:1 的比例。采用 EPI Data 2.1a 统计软件包建立数据库,转化生成 SPSS 11.0 统计软件包数据库。数据分析在 SPSS 11.0 及 PEMS 3.1 统计软件包上实现。

从表 1 协方差分析可见,项目 1 的协变量(入院时的出血量)以及中心与分组交互效应差异均无统计学意义($P > 0.05$),分组间以及中心效应间差异均有统计学意义($P < 0.05$),提示试验组疗效可能优于对照组,且各中心效应趋向不一致;项目 2 的协变量(初次合计关节压痛记分)差异有统计学意义($P < 0.05$),中心效应值差异也有统计学意义($P < 0.05$),提示试验组可能不如对照组疗效。两个项目的中心效应差异均有统计学意义($P < 0.05$),提示两个项目的各中心间效应趋向不一致,且项目 2 还存在分组与中心交互影响,这给分析结果带来比较难于解释的麻烦,为了解决这一问题,我们选择了 Meta 分析分别对两个项目进行了分析。

从表 2 Meta 分析的均数差值及均数差值 95% CI 及图 1 可见,除中心 10 显示差异有统计学意义外,中心 5、6、7 虽然均数差值 > 0 ,但 95% CI 均包括 0,差异无统计学意义,其他中心差值均 < 0 ,95% CI 包括 0,差异也均无统计学意义,项目 1 的各中心异质性检验差异无统计学意义($\chi^2 = 10.52, P = 0.48$),提示中心间结果趋向一致,采用固定效应模型进行合并效应值组间比较,结果表明,合并效应值组间差异有统计学意义($u = 2.31, P = 0.02$),Meta 分析结果与协方差分析结果一致。

对项目 2 行 Meta 分析,结果各中心间效应值(关节压痛个数)异质性检验,差异有统计学意义($\chi^2 = 31.93, P < 0.05$),从表 4 及图 2 可见,中心 1、3、8 的治疗组疗效优于对照组,差异均有统计学意义($P < 0.05$),中心 5 提示治疗组不如对照组,差异也有统计学意义($P < 0.05$),提示各中心间结果趋向不一致,采用随机效应模型进行各中心合并效应

表1 两项多中心临床随机对照试验效应值协方差分析

编号	变异来源	Type III SS	df	均方差	F 值	P 值
项目 1	分组	410.47	1	410.47	6.36	0.012
	中心效应	2291.25	11	208.30	3.23	0.000
	入院时出血量	160.22	1	160.22	2.48	0.116
	分组 * 中心	553.35	11	50.30	0.78	0.660
项目 2	分组	121.80	1	121.80	6.00	0.015
	中心效应	664.28	8	83.04	4.09	0.000
	初次合计关节压痛记分	1145.08	1	1145.08	56.38	0.000
	分组 * 中心	552.41	8	69.05	3.40	0.001

注:SS 为离均差平方和; * 为交互项

表2 项目 1 加权均数差法分析

编号	试验组		对照组		均数差值(95% CI)
	n	$\bar{x} \pm s$	n	$\bar{x} \pm s$	
1	16	8.75 ± 8.00	18	9.06 ± 7.08	-0.31(-5.38~4.76)
2	7	7.43 ± 3.15	2	17.00 ± 21.21	-9.57(-22.98~3.84)
3	14	11.36 ± 11.47	13	12.62 ± 9.84	-1.26(-9.35~6.83)
4	12	4.92 ± 5.60	11	8.09 ± 5.56	-3.17(-7.74~1.40)
5	8	6.63 ± 4.44	7	3.71 ± 6.90	2.92(-2.87~8.71)
6	12	12.58 ± 13.48	16	11.31 ± 11.97	1.27(-8.18~10.72)
7	4	8.50 ± 13.18	6	7.67 ± 8.12	0.83(-12.22~13.88)
8	20	5.30 ± 9.14	15	7.73 ± 8.98	-2.43(-8.50~3.64)
9	17	9.82 ± 6.97	17	10.65 ± 5.85	-0.83(-5.16~3.50)
10	8	3.13 ± 4.91	13	13.54 ± 9.92	-10.41(-17.83~-2.99)
11	16	4.44 ± 5.99	6	6.67 ± 7.28	-2.23(-8.18~3.72)
12	7	1.57 ± 1.51	6	3.33 ± 2.88	-1.76(-4.20~0.682)

值比较,差异无统计学意义($u = 1.39, P > 0.05$),项目 2 的 Meta 分析结果与协方差分析结果不一致(表 4、5,图 2)。

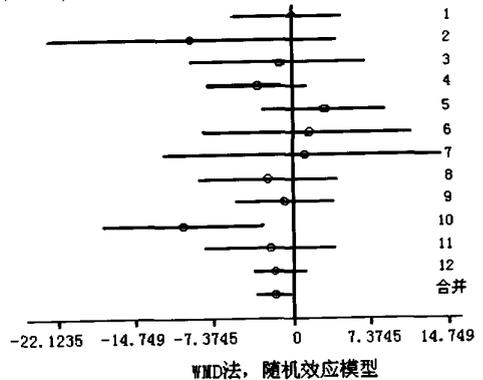


图1 项目 1 各中心效应的 Meta 分析森林图

表3 项目 1 合并统计量及 95% CI

模型	统计量	方差(95% CI)	u 值	P 值
固定效应	-1.75	0.58(-3.25~-0.265)	-2.31	0.021
随机效应	-1.75	0.51(-3.15~-0.350)	-2.45	0.014

注:异质性检验: $\chi^2 = 10.52, P = 0.482$

表4 项目 2 加权均数差法分析

编号	试验组		对照组		均数差值(95% CI)
	n	$\bar{x} \pm s$	n	$\bar{x} \pm s$	
1	20	20.40 ± 8.08	20	24.80 ± 3.85	-4.40(-8.32~-0.48)
2	26	22.70 ± 4.43	26	23.30 ± 6.37	-0.60(-3.58~2.38)
3	27	22.60 ± 4.03	26	25.70 ± 1.64	-3.10(-4.77~-1.43)
4	30	25.20 ± 1.73	28	24.70 ± 1.93	0.50(-0.44~1.44)
5	25	20.10 ± 5.61	25	16.40 ± 5.45	3.70(0.63~6.77)
6	27	22.80 ± 4.83	25	24.50 ± 4.50	-1.70(-4.24~0.84)
7	27	22.00 ± 5.30	31	21.10 ± 5.49	0.90(-1.89~3.69)
8	12	20.70 ± 3.70	15	23.80 ± 3.38	-3.10(-5.78~-0.42)
9	30	23.00 ± 6.80	27	25.50 ± 3.41	-2.50(-5.34~0.34)

表5 项目 2 合并统计量及 95% CI

模型	统计量	方差(95% CI)	z 值	P 值
固定效应	-0.60	0.11(-1.26~0.05)	-1.80	0.07
随机效应	-1.08	0.61(-2.61~0.45)	-1.39	0.17

注:异质性检验: $\chi^2 = 31.9264, P = 0.0001$

Meta分析的森林图

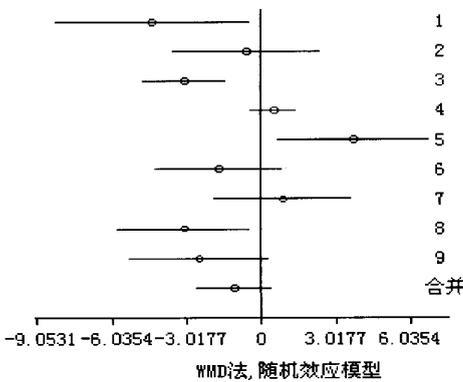


图2 项目 2 各中心效应 Meta 分析森林图

讨论

对广东省中医院国家“十五”科技攻关项目的多中心临床随机对照试验数据处理中,协方差分析结果均提示中心效应差异均有统计学意义($P < 0.05$),其中项目 2 还存在中心与分组的交互效应差异有统计学意义,这给结果的分析带来一定的麻烦和困难,根据 Meta 分析的原理和方法,进一步对两个项目数据分别进行了 Meta 分析,结果项目 1 的 Meta 分析结果中心异质性检验差异无统计学意义,故选择固定效应模型对合并效应值进行组间比较,差异有统计学意义,Meta 分析与协方差分析结果基本一致,因此可以认为,项目 1 的分析结果提示试验组疗效较对照组好;项目 2 异质性检验差异有统计学意义,选择随机效应模型对合并效应值进行组间

比较,差异无统计学意义,项目 2 的 Meta 分析与协方差分析结果不一致,由于项目 2 存在中心与分组交互效应,同时异质性检验差异又有统计学意义,因此,我们认为首先应分析造成项目 2 中心与分组交互效应的可能原因,然后再对项目 2 进行 Meta 分析,并根据 Meta 分析异质性检验的结果选择模型,关于项目 2 中心与分组交互效应的原因还有待进一步的分析。

步的分析。

综上所述,在多中心临床试验的数据分析中,进行中心效应的协方差分析时,如果单纯存在中心效应,而无中心与分组交互效应时,可选择 Meta 分析,根据 Meta 分析异质性检验的结果进一步选择适合的模型分析合并效应值的差异。如果协方差分析结果显示同时存在中心效应、中心与分组交互效应差异均有统计学意义时,在进行 Meta 分析时,首先要分析造成交互效应的可能原因,然后再根据 Meta 分析异质性检验的结果,进一步选择适合的模型,一般协方差分析如果存在交互效应时,通常 Meta 分析的异质性检验差异也有统计学意义,如果协方差分析存在中心与分组交互效应,而 Meta 分析异质性差异无统计学意义,建议 Meta 分析最好选择随机效应模型,因为随机效应模型不仅考虑了研究内变异,同时也考虑了研究间变异,因此其结果比较可靠。

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,2000.12-13.
- 2 Glass GV. Primary, secondary and Meta analysis of research. Educ Res,1976,5:3-8.
- 3 方积乾,陆盈.现代医学统计学.北京:人民卫生出版社,2002.160-209.
- 4 刘关键,吴泰相,康德英. Meta 分析中的统计学过程.中国临床康复,2003,7:538-539.
- 5 康德英,王家良,洪旗,等.临床试验 Meta 分析中的异质性评价研究.华西医学,2000,15:143-145.
- 6 洪楠,林爱华,侯军.SPSS for Windows 统计产品和服务解决方案教程.北京:清华大学出版社,2003.151-163.

(收稿日期:2004-09-16)

(本文编辑:张林东)