

Meta 分析的统计学方法

杨娟, 郑青山

安徽省药物临床评价中心、皖南医学院弋矶山医院, 芜湖 241001, 安徽

摘要 Meta 分析是对具有共同研究目的的相互独立的多个研究结果进行系统合并, 剖析研究间差异特征, 定量综合评价研究结果的统计方法。它在国内外医学文献中的应用日益增多, 其统计学方法也在发展。根据研究的资料类型及统计指标, Meta 分析可选用不同的统计分析方法。本文综述了目前 Meta 分析中应用的比较广泛和成熟的统计学方法。
关键词 Meta 分析; 计量资料; 计数资料; 固定效应模型; 随机效应模型; 异质性检验; 比数比

中图分类号: R969.1

文献标识码: A

文章编号: 1009-2501(2005)11-1309-06

对医学领域某一问题常有许多人研究, 往往出现各研究结果不一致甚至相反的情况, 因此需要对所有结果进行综合。传统的文献综述存在着主观性、片面性、非定量性以及不准确性等弊端, 可能导致不同综述者意见相左。而近年迅速发展的 Meta

分析是对具有共同研究目的的相互独立的多个研究结果进行系统合并, 剖析研究间差异特征, 定量综合评价研究结果的统计方法^[1]。相对于传统的文献综述, Meta 分析扩大样本例数, 提高统计学检验效能、改善对效应量的估计, 因而具有全面、系统、客观、定量的特点。从 1976 年心理学家 Glass 首次正式提出后^[2], Meta 分析在国内外医学文献中的应用日益增多, 其统计学方法也在发展。本文主要介绍目前应用的比较广泛和成熟的 Meta 分析统计学方法。

1 计量资料的 Meta 分析方法

对计量资料进行 Meta 分析可选择均数之差作为统计指标。根据异质性检验 (homogeneity test) 的结果, 选择不同的统计分析模型。

1.1 固定效应模型^[3]

例 1: 为研究氟对女童 II 掌骨皮质厚度的影响, 以 11 个研究女童 II 掌骨皮质厚度为统计指标 (单位 mm) 对表 1 的资料进行 Meta 分析。

表 1 11 个研究女童 II 掌骨皮质厚度 (mm)

研究编号	高氟区			适氟区			合并标准差	效应	效应大小加权合并计算结果		
	n_{1i}	\bar{x}_{1i}	S_{1i}	n_{2i}	\bar{x}_{2i}	S_{2i}			S_{ci}	d_i	w_i
1	25	2.16	0.32	42	2.31	0.33	0.326	-0.460	67	-30.820	14.177
2	55	2.39	0.31	40	2.49	0.32	0.314	-0.318	95	-30.210	9.607
3	46	2.5	0.31	50	2.67	0.35	0.331	-0.514	96	-49.344	25.363
4	28	2.64	0.26	50	2.99	0.55	0.468	-0.748	78	-58.344	43.641
5	45	2.81	0.35	45	2.93	0.36	0.355	-0.338	90	-30.420	10.282
6	52	2.95	0.46	55	3.27	0.37	0.416	-0.769	107	-82.283	63.276
7	36	3.15	0.39	42	3.48	0.48	0.441	-0.748	78	-58.344	43.641
8	36	3.47	0.43	51	3.83	0.54	0.498	-0.723	87	-62.901	45.477
9	43	3.63	0.28	45	3.85	0.44	0.371	-0.593	88	-52.184	30.945
10	42	3.81	0.31	45	4.16	0.42	0.371	-0.943	87	-82.041	77.365
11	41	3.99	0.55	27	4.18	0.41	0.500	-0.380	68	-25.840	9.819
合计									941	-562.731	373.593

2005-10-25 收稿 2005-11-14 修回

安徽省优秀青年科技基金项目 (04043052)

杨娟, 女, 硕士研究生, 研究方向: 临床药理学和生物统计学。

Tel: (0) 13855367615 E-mail: janyang5006@163.com

郑青山, 通讯作者, 男, 博士, 教授, 研究员, 研究方向: 临床药理学和生物统计学。

Tel: 0553-5738350 E-mail: drugchina@21cn.com

设 k 项研究报告中, 第 i 项研究的处理组和对照组的样本量分别为 n_{1i} 和 n_{2i} , 均数分别为 \bar{x}_{1i} 和 \bar{x}_{2i} , 标准差分别为 S_{1i} 和 S_{2i} 。分析如下:

1. 计算每个研究的标准化均数之差及两组的合

并方差

$$\text{标准化均数之差: } d_i = \frac{\bar{x}_{1i} - \bar{x}_{2i}}{s_{ci}}, i = 1, 2, 3 \dots k$$

两组的合并方差:

$$s_{ci} = \sqrt{\frac{(n_{1i} - 1) s_{1i}^2 + (n_{2i} - 1) s_{2i}^2}{n_{1i} + n_{2i} - 2}}$$

2. 计算权重系数和加权均数

$$\text{权重系数: } w_i = n_{1i} + n_{2i}$$

$$\text{加权均数: } \bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

3. 异质性检验

$$\text{加权方差的估计值: } S_d^2 = \frac{\sum w_i d_i^2 - \bar{d}^2 \sum w_i}{\sum w_i}$$

$$\text{异质性校正因子: } S_e^2 = \frac{4k}{\sum w_i} \left(1 + \frac{\bar{d}^2}{8}\right)$$

$$\text{检验统计量: } \chi^2 = \frac{k S_d^2}{S_e^2}$$

此检验的零假设为各研究的总体效应相同。统计量服从自由度为 $k - 1$ 的 χ^2 分布。按一定的显著性水平 做出统计推断,若 $P >$, 则支持固定效应模型的假定。

4. 总体平均效应的 95 %可信区间 (95 %CI)

固定效应模型分析时, \bar{d} 的标准误为:

$$S_{\bar{d}} = \frac{S_e}{\sqrt{k}}$$

总体平均效应的 95 %可信区间为: $\bar{d} \pm 1.96 S_{\bar{d}}$

例 1 中:

$$\bar{d} = \frac{-562.731}{941} = -0.598$$

$$S_d^2 = \frac{373.593 - (-0.598)^2 \times 941}{941} = 0.0394$$

$$S_e^2 = \frac{4 \times 11}{981} \left(1 + \frac{(-0.541)^2}{8}\right) = 0.0488$$

$$\chi^2 = \frac{11 \times 0.0394}{0.0488} = 8.88$$

$= 11 - 1 = 10$, 查 χ^2 分布表, $P > 0.05$, 认为 11 个研究具有同质性, 可采用固定效应模型进行分析。 \bar{d} 的标准误 $S_{\bar{d}} = 0.066$, 总体平均效应的 95 %CI 为: $-0.598 \pm 1.96 \times 0.066$, 即 $-0.73 \sim -0.47$, 未包括 0, 综合 11 个研究, 可认为高氟区女童 II 掌骨皮质平均厚度约比适氟区低 0.47 至 0.73 倍标准差。

1.2 随机效应模型^[1] 上例中, 若异质性检验 $P < 0.05$, 拒绝零假设, 则采用随机效应模型。主要的统

计分析方法步骤如下:

1. 计算每个研究的标准化均数之差及两组的合并方差

$$\text{标准化均数之差: } d_i = \frac{\bar{x}_{1i} - \bar{x}_{2i}}{s_{ci}}, i = 1, 2, 3 \dots k$$

两组的合并方差:

$$s_{ci} = \sqrt{\frac{(n_{1i} - 1) s_{1i}^2 + (n_{2i} - 1) s_{2i}^2}{n_{1i} + n_{2i} - 2}}$$

2. 计算权重系数和加权均数

$$\text{权重系数: } w_i = n_{1i} + n_{2i}$$

$$\text{加权均数: } \bar{d} = \frac{\sum w_i d_i}{\sum w_i}$$

3. 异质性检验

$$\text{加权方差的估计值: } S_d^2 = \frac{\sum w_i d_i^2 - \bar{d}^2 \sum w_i}{\sum w_i}$$

$$\text{异质性校正因子: } S_e^2 = \frac{4k}{\sum w_i} \left(1 + \frac{\bar{d}^2}{8}\right)$$

$$\text{检验统计量: } \chi^2 = \frac{k S_d^2}{S_e^2}$$

若 $S_d^2 > S_e^2$, 则校正后的标准误为: $S = \sqrt{S_d^2 - S_e^2}$

效应合并值的 95 %CI 为: $\bar{d} \pm 1.96 S$

若 $S_d^2 < S_e^2$, 则随机效应模型退化为固定效应模型。在区间估计时, 随机效应模型比固定效应模型估计更精确(区间宽度更窄)。

2 计数资料的 Meta 分析方法

对于计数资料, 主要讨论四格表资料的 Meta 分析。能够形成四格表资料的研究方法最常见的有随机化临床试验、病例-对照研究、队列研究和诊断试验评价等。这些研究的数据基本格式见表 2。另外等级资料转换成二分类资料后也可用计数资料的 Meta 分析方法进行合并。

2.1 固定效应模型 对于计数资料而言, 适应于固定效应模型的 Meta 分析方法有 Mantel-Haenszel 法(简称 M-H 法)、Peto 法、Fleiss 法^[4]以及 General variance-based 法。Mantel-Haenszel 法是利用分层分析的原理, 将每一层作为一个独立研究, 计算综合的 OR 值并检验^[5]。在此结合实例介绍改良的 M-H 法即 Peto 法并介绍 Fleiss 法和总方差分析法和总方差分析法(General variance-based 法)的分析步骤。为了统一方法中的参数, 特规定一些参数意义如表 2:

表 2 四格表资料的基本格式

组别	阳性	阴性	合计
处理组	a_i	b_i	n_{1i}
对照组	c_i	d_i	n_{2i}
合计	m_{1i}	m_{2i}	N_i

例 2:激素预防新生儿肺透明膜病的 14 个研究资料如表 3,可见各个独立研究的结果不尽相同,在这种情况下,做一个 Meta 分析是很有价值的,现分别用这两种方法进行统计合并。

表 3 激素预防新生儿肺透明膜病的研究

研究编号	激素组		对照组		OR	95%可信区间
	发病人数	例数	发病人数	例数		
1	35	533	60	536	0.56	0.36—0.87
2	1	61	5	54	0.16	0.02—1.42
3	16	133	23	140	0.7	0.35—1.39
4	3	67	7	60	0.35	0.09—1.42
5	9	57	10	71	1.14	0.43—3.03
6	3	64	13	59	0.17	0.05—0.63
7	1	71	7	75	0.14	0.02—1.17
8	4	81	11	63	0.25	0.08—0.83
9	32	372	33	371	0.96	0.58—1.60
10	5	49	2	29	1.53	0.28—8.45
11	7	121	13	124	0.52	0.20—1.35
12	1	24	1	22	0.91	0.05—15.49
13	5	36	6	37	0.83	0.23—3.01
14	6	96	8	92	0.70	0.23—2.10

2.1.1 Peto 法 在 Cochrane 协作网提供的 Meta 分析软件 RevMan 4.2 中采用的即是 Peto 法。它适用对以比数比 OR(odds ratio) 为效应指标的多个研究

进行合并,是固定效应模型的经典方法。此方法的分析过程如下(计算合并所需要的中间值见表 4):

表 4 激素预防新生儿肺透明膜病 Peto 法计算合并所需要的中间值

研究编号	O_i	E_i	V_i	$O_i - E_i$	$(O_i - E_i)^2 / V_i$
1	35	47.367	21.659	- 12.367	7.061
2	1	3.183	1.429	- 2.183	3.335
3	16	19	8.382	- 3.00	1.074
4	3	5.276	2.314	- 2.276	2.239
5	9	8.461	4.028	0.539	0.072
6	3	8.325	3.502	- 5.325	8.097
7	1	3.89	1.902	- 2.89	4.391
8	4	8.438	3.33	- 4.438	5.915
9	32	32.544	14.848	- 0.544	0.02
10	5	4.397	1.508	0.603	0.241
11	7	9.878	4.61	- 2.878	1.797
12	1	1.043	0.488	- 0.043	0.004
13	5	5.425	2.368	- 0.425	0.076
14	6	7.149	3.255	- 1.149	0.406
合计	35		73.623	- 36.376	34.728

1. 计算每个研究的 OR_i : $OR_i = \frac{a_i d_i}{b_i c_i}$
2. 计算每个研究中某事件发生数的期望值 E_i :

$$E_i = \frac{m_{1i} n_{1i}}{N_i}$$

3. 计算每个研究中某事件发生数的方差 V_i :

$$V_i = \frac{m_{1i} n_{1i} m_{2i} n_{2i}}{N_i^2 (N_i - 1)}$$

4. 异质性检验的统计量

$$Q = \sum \frac{(O_i - E_i)^2}{V_i} - \frac{(\sum(O_i - E_i))^2}{\sum V_i} = k - 1$$

其中 O_i 为第 i 个研究处理组阳性结果的实际值。此检验的零假设为各研究的总体效应相同。统计量服从自由度为 $= k - 1$ 的 χ^2 分布。按一定的显著性水平 α 做出统计推断,若 $P > \alpha$,表明可进行合并分析。

5. 计算合并的 OR 值:

$$OR_p = \exp\left(\frac{\sum(O_i - E_i)}{\sum V_i}\right)$$

6. 计算 OR_p 的 95 %CI:

$$\exp\left[\frac{\sum(O_i - E_i) \pm 1.96 \sqrt{\sum V_i}}{\sum V_i}\right]$$

7. OR_p 假设检验的统计量

$$\chi^2 = \frac{(\sum(O_i - E_i))^2}{\sum V_i} = 1$$

例 2 中, $Q = 34.7288 - \frac{(-36.376)^2}{73.623} = 16.75$

$= k - 1 = 14 - 1 = 13$,查表 $\chi^2(0.05, 13) = 22.$

36, 故 $P > 0.05$,表明可进行合并分析。

$$OR_p = \exp\left(\frac{\sum(O_i - E_i)}{\sum V_i}\right)$$

$$= \exp\left(\frac{-36.376}{73.623}\right) = 0.61$$

OR_p 的 95 %CI:

$$\exp\left[\frac{\sum(O_i - E_i) \pm 1.96 \sqrt{\sum V_i}}{\sum V_i}\right] = \exp$$

$$\left(\frac{-36.376 \pm 1.96 \sqrt{73.623}}{73.623}\right) = (0.49 \sim 0.77)$$

OR_p 假设检验的统计量:

$$\chi^2 = \frac{(\sum(O_i - E_i))^2}{\sum V_i} = \frac{(-36.376)^2}{73.623} = 17.97$$

$= 1$, $P < 0.01$,说明使用激素的新生儿肺透明膜病的发病率比不使用激素的新生儿低。

2.1.2 Fleiss 法 此法是 Fleiss 提出的以率 P_i 直接计算 OR ,可用于临床试验研究和队列研究。方法

如下:

1. 计算各研究的 OR_i : $OR_i = \frac{P_{1i}/(1 - P_{1i})}{P_{2i}/(1 - P_{2i})}$

2. 计算每个研究的方差和权重:

$$V_i = \frac{1}{n_{1i} P_{1i} (1 - P_{1i})} + \frac{1}{n_{2i} P_{2i} (1 - P_{2i})}$$

$$W_i = \frac{1}{V_i}$$

3. 异质性检验的统计量:

$$Q = \sum W_i (1 - \ln(OR_i) - \ln(OR_p))^2 = \sum W_i \ln(OR_i)^2 - \frac{(\sum W_i \ln(OR_i))^2}{\sum W_i}$$

此检验的零假设为各研究的总体效应相同。统计量服从自由度为 $= k - 1$ 的 χ^2 分布。按一定的显著性水平 α 做出统计推断,若 $P > \alpha$,表明可进行合并分析。

4. 计算合并的 OR 值:

$$OR_p = \exp\left(\frac{\sum W_i \ln(OR_i)}{\sum W_i}\right)$$

5. 计算 OR_p 的 95 %CI:

$$\exp(\ln(OR_p) \pm 1.96 / \sqrt{\sum W_i})$$

6. OR_p 假设检验的统计量:

$$\chi^2 = \frac{(\sum W_i \ln(OR_i))^2}{\sum W_i}$$

自由度为 $= 1$,查 χ^2 界值表,若 $P < 0.05$,说明处理组和对照组差别有统计学意义。

2.1.3 总方差分析法 (General variance-based 法)^[11]

该方法适用于以率比或率差为测定指标的队列研究及临床试验结果的合并。多用于分析危险率差 (RD)和危险率比 (RR)。在此仅介绍以 RD 为统计指标的分析过程。仍以表 2 中参数说明主要计算过程如下:

1. 计算每个研究的危险率差: $RD_i = \frac{a_i}{n_{1i}} - \frac{c_i}{n_{2i}}$

2. 计算每个研究的方差和权重:

$$Var(RD_i) = \frac{m_{1i} m_{2i}}{n_{1i} n_{2i} N_i}$$

$$W_i = 1 / Var(RD_i)$$

3. 求合并 RD 值: $RD_G = \frac{\sum(W_i RD_i)}{\sum W_i}$

4. 求 RD_G 的 95 %CI: $RD_G \pm 1.96 / \sqrt{\sum W_i}$

5. 异质性检验统计量:

$$Q = \sum W_i (RD_i - RD_G)^2$$

检验统计量 Q 服从自由度为 $= k - 1$ 的 χ^2 分

布。按一定的显著性水平 做出统计推断,若 $P >$,表明可进行合并分析。

6. RD_G 的假设检验:

$$Z^2 = \frac{RD_G^2}{Var(RD_G)}$$

$$Var(RD_G) = 1 / \sum W_i$$

自由度为 $k - 1$,查 Z^2 界值表,若 $P < 0.05$,说明处理组和对照组差别有统计学意义。

2.2 随机效应模型 计数资料在进行 Meta 分析时,若异质性检验拒绝零假设,应采用随机效应模型。该模型由 DerSimonian 和 Laird 于 1986 年提出^[1,6],是基于随机效应模型的经典方法。

1. 计算各个研究的 OR_i : $OR_i = \frac{a_i d_i}{b_i c_i}$

2. 用 $M-H$ 法估计 OR_{MH} 值:

$$OR_{MH} = \frac{\sum \frac{b_i c_i}{N_i} OR_i}{\sum \frac{b_i c_i}{N_i}}$$

3. 计算异质性检验的 Q 值:

$$Q = \sum \frac{b_i c_i}{N_i} [\ln(OR_i) - \ln(OR_{MH})]^2$$

$= k - 1$,查 Z^2 界值表,若 $P < 0.05$, k 个研究间异质性检验拒绝零假设,不能应用固定效应模型,而要用随机效应模型的计算过程。

4. 计算校正因子 D 若 $Q < k - 1$,则 $D = 0$;若 $Q > k - 1$,则按下式计算 D :

$$D = \frac{[Q - (k - 1)] \sum \frac{b_i c_i}{N_i}}{(\sum \frac{b_i c_i}{N_i})^2 - \sum (\frac{b_i c_i}{N_i})^2}$$

5. 用 DerSimonian-Laird 法计算权重:

$$W_i = \frac{1}{D + (1/\frac{b_i c_i}{N_i})}$$

6. 用 DerSimonian-Laird 法计算合并的 OR 值:

$$OR_{DL} = \exp\left(\frac{\sum W_i \ln(OR_i)}{\sum W_i}\right)$$

7. 计算 95%CI:

$$\exp(\ln(OR_{DL}) \pm 1.96 / \sqrt{\sum W_i})$$

3 检验统计量的综合分析^[7]

一般对单个研究所作的统计分析,可以是 t 检验、 u 检验,或得出相应的 P 值。Meta 分析的综合检验分析方法是对同类单个独立研究的检验统计量

进行合并分析,求一综合检验统计量,用以确定有无统计学意义。常用方法有:

3.1 合并 P 值的 Fisher 法 设 P_i 为按第 i 个试验的假设检验结果算得的单侧检验概率, k 项独立研究的综合检验统计量为:

$$X^2 = \sum (-2 \ln P_i)$$

服从自由度为 $2k$ 的 X^2 分布,由此可得出相应的合并 P 值并进行统计推断。

3.2 Stouffer 法 将各独立研究的单侧概率 P 值转换为相应的标准正态分布离差值,记为 u_i 。则由下式计算的 k 个独立研究的综合检验统计量服从标准正态分布。

$$u = \frac{\sum u_i}{\sqrt{k}}$$

3.3 Winner 法 记 t_i 为各独立研究假设检验的统计量 t 值, v_i 为各研究对应的自由度,则下式计算的统计量近似服从标准正态分布。

$$u = \frac{\sum t_i}{\sqrt{\sum v_i / (v_i - 2)}}$$

3.4 X^2 检验的综合分析法 记 \bar{x}_i^2 为各独立研究假设检验的统计量,则 k 个研究的 X^2 值之和服从自由度为 k 的 X^2 分布。

$$X^2 = \sum x_i^2$$

除以上方法外,近些年国内学者也构建了一些新的 Meta 分析统计模型。如有人提出 Meta 分析的多水平模型,除可估计平均效应尺度及其可信区间外,可分析与研究结果差异有关的影响因素^[8]。有人构建了 Meta 分析的 Bayes 系统可交换模型,采用 WinBUGS 软件由 Gibbs 抽样的 MCMC 模拟技术估计参数^[9]。还有人建立了模糊数学 Meta 分析法用于药物疗效和药物经济学评价,以得出模糊性的定量结论^[10]。此外,我们也可以将各独立研究的统计量看作反应变量,用方差分析、多元回归、logistic 模型、对数线性模型和 Cox 模型等进行合并统计量及齐性检验的多因素分析。

总之,根据研究的资料类型及统计指标,Meta 分析可选用不同的统计分析方法,定量综合分析多个研究结果,得出较为合理的结论,为临床用药提供依据,为防治策略与研究方案的制定等提供更可靠的依据。现在,美国已有新药申报资料应用 Meta 分析支持增加适应证以及进行亚组分析^[11],欧盟则要求用几个设计较好的临床试验的 Meta 分析结果证明新药的有效性^[12]。恰当的应用 Meta 分析将是对

世界范围内的医学资源的节约和有效的利用。

参考文献

- 1 孙瑞元,郑青山,主编.数学药理学新论[M].人民卫生出版社,2004:631 - 66
- 2 Glass GV. Primary, secondary, and meta-analysis of research [J]. Education Research,1976;6:3
- 3 陈长生,徐勇勇.第十四讲,如何进行Meta分析[J].中华预防医学杂志,2003;37:138 - 40
- 4 Fleiss JL, Grss AJ. Meta-analyses in epidemiology with special reference to studies of association between exposure to environmental tobacco smoke and lung cancer; a critique [J]. J Cho Epi,1991;2:127
- 5 富振英. Meta-analysis 方法介绍 [J]. 卫生研究,1995;24:380

- 6 DerSimonian R,Laird N. Meta-analysis in clinical trials. Contr Clin Trial,1986;7:177 - 188
- 7 刘玉秀,洪立基. 新药临床研究设计与统计分析[J]. 南京大学出版社,1999;273 - 4
- 8 李小松,刘巧兰,倪宗瓚. 多水平统计模型在 Meta 分析中的应用研究[J]. 中国卫生统计,1999;16:133 - 5
- 9 郭秀娥,徐勇勇,张高魁. 临床试验四格表资料 Meta 分析的 Bayes 方法[J]. 第四军医大学学报,2000;21:679 - 81
- 10 朱含涌等. 模糊数学 Meta 分析的理论与方法研究[J]. 药物流行病学杂志,2002;11:311 - 4
- 11 Anello C. Emerging and recurrent issues in drug development [J]. Statistics Med,1999;18:2301 - 9
- 12 Willem K,Scholten,MSc,Pharm,MPA. Introducing a system for assessment of quality, safety, and efficacy of herbal medicinal products in the Netherlands[J]. Drug inform J,2001;2:461 - 8

Statistical methods of Meta-analysis

YANG Juan,ZHENG Qing-shan

Anhui Provincial Center for Drug Clinical Evaluation & Yijishan Hospital of Wannan Medical College, Wuhu 241001, Anhui, China

ABSTRACT Meta-analysis is a statistical method to systematically combine and evaluate several relevant trials. It is applied generally in medical literatures, and its statistical methods are developing. The aim of this paper is to introduce several widely-used statistical methods of

Meta-analysis.

KEY WORDS Meta-analysis; measurement data; enumeration data; fixed effect model; random effect model; test for heterogeneity; odds ratio