

Meta 分析

张小年

[关键词] 医学 ; 科学研究 ; 分析方法
中图分类号 : R195.4 文献标识码 : B 文章编号 : 1006-9771 (2005) 08-0686-02
[本文著录格式] 张小年 . Meta 分析 [J] . 中国康复理论与实践 , 2005 , 11 (8) : 686 — 687 .

Meta 分析 (Meta-analysis) 常译为汇总分析或荟萃分析 , 是系统评价 (systematic review , SR) 中经常用到的统计方法。正是 Meta 分析被引入医学科学 , 才解决了综合几个研究结果得出一个总体定量数据的问题 , 使 SR 具有较强的科学性和可重复性。运用 Meta 分析方法 , 可避免各个研究组之间可能由于取样来自不同的总体而造成的差异 , 也可根据研究的样本量为其结论赋予不同的权重 , 而非简单的合并。因此 , Meta 分析最大的优点是通过增大样本量提高结论的可信度 , 降低研究结果的不一致性。Meta 分析可用于临床试验、诊断试验和流行病学研究的 SR。但 Meta 分析并不是万能的 , 对一些经过大样本、多中心的临床试验已得到明确结论的 , 不必做 Meta 分析 ; 而对那些设计或执行质量很差的研究 , 如有偏倚的或没有意义的资料 , 也不能寄希望于通过 Meta 分析就能得到可信的结论。国际上对评估证据强度作了分级 (见表 1) , 也制订了 Meta 分析的最低标准^[1-4] , 即 , 随机对照研究 (randomized controlled trial , RCT) 质量特征 : ①良好设计 ; ②良好质控 ; ③控制偏差 ; ④有效样本 ; ⑤确切终点 ; Meta 分析质量特征 : ①前瞻设计 ; ②准确终点 ; ③同质对象 ; ④涵盖全部研究和病例 , 以保证其评估质量。

表 1 评估证据强度分级的建议^[1]

分级	建议
A 级	方法最可靠 , 多个 RCT 研究结果一致
	1. 效果明确 , 效益确实大于 (或不大于) 风险
B 级	2. 效果不能肯定效益是否大于风险
	方法可靠 , 多个 RCT 研究结果不完全一致
C 级	1. 效果明确 , 效益确实大于 (或不大于) 风险
	2. 效果可疑 , 不能肯定效益是否大于风险
	方法论证强度低 , 观察性研究
	1. 效果明确 , 效益确实大于 (或不大于) 风险
	2. 效果可疑 , 不能肯定效益是否大于风险

1 常用测量指标

1.1 计量资料 通常情况下 , 每一项研究的结果可用四格表显示 (见表 2) , 再据此计算出选用的评价指标^[5,6]。

表 2 四格表

	试验组 (暴露组)	对照组	合计
发生事件 (发病或死亡)	a	b	a + b
未发生事件	c	d	c + d
合计	a + c	b + d	a + b + c + d

作者单位 : 100068 北京市 , 北京博爱医院 ; 2. 100068 北京市 , 首都医科大学康复医学院。作者简介 : 张小年 (1978-) , 男 , 浙江台州市人 , 医学博士 , 主要研究方向 : 神经康复及睡眠神经生理。

相对危险度 (risk ratio , RR) : $RR = [a / (a + c)] / [b / (b + d)]$, 指某因素的暴露组发生结局事件的危险度与对照组发生结局事件的危险度的比值。RR = 1.0 表示两组危险度相等 , 考察因素对该病没有影响 (或治疗无效)。RR < 1 且有显著性差异时 , 说明考察因素是保护因素 (或是有效的治疗方法) ; RR > 1 且有显著性差异时 , 说明考察因素是危险因素 (或治疗反而有害)。

相对危险度减少 (relative risk reduction , RRR) : $RRR = 1 - RR$, 指与对照组相比 , 治疗组结局事件发生减少的百分比 , 即治疗组经治疗后 , 结局事件发生的 RR 值下降的水平。通常 RRR 为 25 % ~ 50 % 或更大时才有临床意义。

比数比 (odds ratio , OR) : $OR = (a / b) / (c / d) = a d / b c$, 指发生结局事件组中暴露于某因素的比数与未发生结局事件组中暴露于某因素的比数的比值。RR 与 OR 似乎差别不大 , 很多文献中也常常将两者等同起来 , 这其实是临床试验中常见的错误之一。RR 一般用于前瞻性研究 , OR 一般用于回顾性研究。如果发病率低 ($a \approx a + c$) , OR 近似等于 RR , 随着发病率的增加 , 两者之间的差异也增大。

危险差 (risk difference , RD) : $RD = [a / (a + c)] - [b / (b + d)]$, 指治疗组与对照组结局事件危险度的绝对差值。RD < 0 且有显著性差异时 , 表明治疗有效 ; RD > 0 且有显著性差异时 , 表明治疗反而有害。RD 不同于 OR 和 RR , 是一个绝对测量指标 , 较适用于将结果应用于某一特定的医疗情况中。此外 , 用 RD 进行 Meta 分析时 , 即使试验的发生事件数 (a 或 c) 为 0 , 该项试验的结果仍可纳入分析 , 而用 OR 或 RR 分析时则不能。

需要治疗的例数 (number needed to treat , NNT) : $NNT = 1 / RD$, 指为挽救一个患者免于发生结局事件 , 需要治疗具有发生此类危险性患者的总人数 (number needed to treat to benefit , NNTb) , 常用来表示疗效结果。NNTb 很容易用 RD 计算得到 , 而用 OR 或 RR 计算则相对较难。

1.2 等级资料 加权均数差 (weighted mean difference) : 如果不同的研究均采用统一的标准测量结局 , 就可计算加权均数差 , 其优点在于以简明易懂的自然单位对结果进行概括。

标准化均差 (standardised mean differences) : 偶尔对那些定义相同而采用了不同指标测量结局的研究进行结果的合并时 , 可使用标准化均差 (例如对生活质量的测量量表)。由于这种方法可以将极不相同的研究结果进行合并 , 因此使用这种方法要特别谨慎 , 尤其要清楚合并结果是否有意义 , 对结果的解释也要慎重。

2 统计模型和方法的选择

2.1 评价资料的统计齐性 各项研究之间的差异可归因于两个来源。首先是随机抽样误差。即使两个研究的总体效应完全相同,即可认为来自同一总体,因为随机抽样误差也可得到不同的结果,但与实际效应(真值)相差不会很大。这种差异称为研究组内方差。第二种误差是由于对象来自不同的总体,即使干预的措施和其他情况都一样,其实际效应也可能不相同。这种差异称为随机效应,它描述了研究组间的变化。检验研究组间是否齐性可应用 χ^2 检验。进行 Meta 分析前必须进行齐性检验,但齐性检验的把握度很低,齐性检验认为研究间存在齐性并不表明各个体研究有相同的效果,只表示统计学上可以假设各个体研究样本来自同一总体。

如果研究之间的差异很大,存在异质性,这时直接计算合并的效应大小很危险,需要仔细检查研究设计和其他可能导致异质性的原因^[7]。主要可以从以下几个方面考虑各研究的差异:设计质量、干预措施、测量结果的时点与测量方法、分析使用的统计模型及方法、个体效应量、纳入和排除标准等。分析发现异质性的来源后,就可以根据原因作处理:①亚组分析:可以将各研究按设计方案、研究质量等分为几个亚组,分别合并及评价;②敏感性分析:将所有研究纳入分析,得出结果,再将异常的研究排除,重新分析,比较前后两次的结果,探讨被排除的研究对结果的影响,进一步还可以进行 Meta 回归分析;③采用随机效应模型;④Meta 回归及混合模型:利用回归模型控制混杂。

2.2 模型及方法的选择 Meta 分析所用的统计方法分为固定效应模型(fixed effects model)和随机效应模型(random effects model)。固定效应模型的统计方法假设个体研究的方差齐性,所以,其效应大小综合估计的方差成分只包括了各个体研究内的方差。而随机效应模型并不假设个体研究方差的齐性,其效应大小综合估计的方差成分既包括了各个体研究内的方差,也包括了各个体研究之间的方差。所以,后者所得到的结果 95% 可信区间较大,结果也更保守。在 Meta 分析中,如各个体研究的资料为非齐性时,应该选择随机效应模型的统计方法。

两种模型中常用的统计方法见表 3。

表 3 固定效应模型与随机效应模型统计方法

模型假设	统计方法	效果测量形式
固定效应模型 (方差齐性)	Mantel-Haenszel 法	比(OR 或 RR)
	Peto 法	比数比(OR)
	General Variance-Based 法	比(所有类型)、差值(RD、平均数差值等)和回归系数
随机效应模型	DerSimonian-Laird 法	比(所有类型)、差值(RD、平均数差值等)

Mantel-Haenszel 法:本法具有理想的统计特征,比较符合数学理论,但需要每个研究有完整的四格表,如果某个研究无法完成四格表,这个研究必须被剔除。但 Meta 分析的资料往往来自文献报道而无法获得原始数据,因此限制了本法的应用。另外,这种方法不考虑混杂因素,在严格的随机对照研究和病例对照研究中,混杂因素已经被配对^[8],因此使用这种方法问题不大,否则就可能导致很大的误差。

Peto 法:本法是 Mantel-Haenszel 法的改良方法,比较简便易行,而且解决了 Mantel-Haenszel 法 χ^2 和 OR 有时不一致的情况。但该方法在四格表很不平衡或各个纳入研究的 OR 值

都远离 1(>5 或 <2)时会产生很大的偏性,故只能用于 OR 值的处理,很少用于非实验性研究,而且也同样存在 Mantel-Haenszel 法的缺陷。

General Variance-Based 法:可对多种指标进行合并分析,包括比、差值和回归系数等。如果研究只提供了最终结果(如 RR 和 95%CI)而无法获得原始四格表时,仍可以利用近似公式计算。

DerSimonian-Laird 法:(略)。

以上各统计方法都可利用统计软件(如 Rev Man)完成。

3 分析结果显示

利用 Rev Man 软件可以方便地将分析结果以图表形式直观地显示。经常使用的是水平条图,又称森林图^[9]。图中显示了每个研究的作者或名称、试验组的结局事件出现例数(n)和总人数(N)、对照组的结局事件出现例数(n)和总人数(N)、观察值(OR 或 RR 等)的图示及具体数值。对于结局事件的定义一般在图表标题框中说明。

每个研究的 OR(RR)值都显示为实心的方块,方块的大小代表了该研究样本量的多少。穿过其中心的水平线代表了该值的置信区间(CI)。Meta 分析的结果列于参与该分析的几个研究之后,用菱形表示,其水平宽度代表了置信区间的范围。中央的竖线代表 OR 值为 1,即干预无效。OR 值的置信区间如果穿过这条竖线,就意味着 CI 中包含 1,本研究无显著性差异。如果置信区间完全落在竖线左侧,说明试验组发生的结局事件的例数少于对照组;落在竖线右侧则意义相反。在每个 Meta 分析结果的下方都会给出各试验间异质性检验的 P 值,一般 P>0.05 就认为试验间方差齐性。此外还有总效应测定。如果 Meta 分析分几个亚组进行,每个亚组后分别列出 Meta 分析的结果,最后列出所有研究的总结果。

[参考文献]

[1]Gorden HG, Deborah JC, David LS, et al. Grades of recommendation for antithrombotic agents[J].Chest,1998,114:441—444.

[2]Gorden HG, Jack S, Deborah JC, et al. Users' guides to the medical literature: XVI. How to use a treatment recommendation[J].JAMA,1999,281:1836—1843.

[3]李强.循证医学——临床证据的产生、评价与利用[M].北京:科学出版社,2001.

[4]王吉耀.循证医学与临床实践[M].北京:科学出版社,2002.

[5]曹家琪.临床医学研究方法学[M].北京:北京医科大学与中国协和医科大学联合出版社,1993.

[6]徐德忠.临床科研方法与实例评价[M].西安:陕西科学技术出版社,1993.

[7]Esther MJ, Steultjens EM, Dekker J, et al. Occupational therapy for stroke patients: a systematic review[J].Stroke,2003,34(3):676—687.

[8]Giibertson L, Langhorne P, Walker A, et al. Domiciliary occupational therapy for patients with stroke discharged from hospital: a randomised trial[J].BMJ,2000,320:603—606.

[9]Langhorne P, Greener J. Systematic reviews in rehabilitation for stroke: issues and approaches to addressing them[J].Clin Rehabil,2002,16:69—74.

(收稿日期:2004-10-31)