

# Meta分析的 统计学过程

刘关键

四川大学华西医院  
循证医学与临床流行病学中心

- “Meta”一词源于希腊文，意为“more comprehensive”。即“更广泛、更全面”
- 上个世纪60年代开始，在医学文献中，陆续出现了对多个独立研究的统计量进行合并的报道。

- 英国心理学家G. V. Glass, 1976年首先将这种多个同类研究的统计量合并方法称为“Meta-Analysis”。
- Glass最早在教育学研究中使用了Meta分析，而该方法现已广泛应用于医学和健康领域，尤其是针对疾病的诊断、治疗、预防和病因等问题的综合评价。

- 上世纪80年代末，我国医学杂志较多的介绍了该方法。其中文译名有荟萃分析，二次分析、汇总分析、集成分析等。
- 无论何种中文译名都有不足之处。因此，很多学者建议仍然使用“Meta分析”这一名称。

## 一、Meta分析的定义

- **Meta-Analysis is a systematic review that uses quantitative methods to summarize the results.**
- Meta分析是运用定量方法去概括（总结）多个研究结果的系统评价。  
《Evidence-Based Medicine》--David Sackett等，第247页的定义

- **Meta-Analysis is statistical technique for assembling the results of several studies in a review into a single numerical estimate.**
- **Meta分析是文献评价中，将若干个研究结果合并成一个单独数字估计的统计学方法。**
- **《The Cochrane Library》第3页的定义。**

## 二、Meta分析的统计目的

### 实例一

#### 七个阿斯匹林预防心肌梗死的研究资料

(Fleiss JL, Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. J Clin Epidemiol. 1991;44(2):127-39.)

K 个研究	阿斯匹林		安慰剂		合计	OR	95%CI	
	死亡数	未死亡数	死亡数	未死亡数			下限	上限
	$a_i$	$b_i$	$c_i$	$d_i$				
1	49	566	67	557	1239	0.720	0.489	1.059
2	44	714	64	707	1529	0.681	0.457	1.013
3	102	730	126	724	1682	0.803	0.606	1.063
4	32	285	38	271	626	0.801	0.486	1.319
5	85	725	52	354	1216	0.798	0.553	1.153
6	246	2021	219	2038	4524	1.133	0.935	1.373
7	1570	7017	1720	6880	17187	0.895	0.829	0.966
合计	2128	12058	2286	11531	28003			

表中 $a_i$ 、 $b_i$ 、 $c_i$ 、 $d_i$ 为各研究四格表数， $N_i$ 为各个研究的样本例数。

### 实例二

#### 女童掌骨Ⅱ皮质厚度的11个研究

K 个研究	高氟区			适氟区			P 值
	$n_1$	$\bar{X}_1$	$s_1$	$n_2$	$\bar{X}_2$	$s_2$	
1	26	2.26	0.32	42	2.33	0.33	$P > 0.05$
2	55	2.39	0.31	40	2.49	0.32	$P > 0.05$
3	46	2.50	0.30	50	2.67	0.35	$P < 0.05$
4	45	2.64	0.26	50	2.90	0.45	$P < 0.05$
5	45	2.81	0.35	45	2.93	0.36	$P > 0.05$
6	52	2.95	0.46	55	3.27	0.37	$P < 0.05$
7	46	3.15	0.39	42	3.48	0.48	$P < 0.05$
8	45	3.47	0.46	51	3.73	0.54	$P < 0.05$
9	45	3.63	0.38	45	3.81	0.40	$P < 0.05$
10	42	3.81	0.41	45	4.16	0.42	$P < 0.05$
11	44	3.99	0.56	25	4.18	0.41	$P > 0.05$
合计	491			490			

方积乾.医学统计学与电脑实验.第二版.上海科学技术出版社,2001,349-350.

### 传统文献综述的特点

- 在医学研究中，传统的文献综述在处理同一问题的多个结果报道时，通常是平等（等权重方法）对待每个研究结果而得出结论。这种文献综述主要是以某类结果的文献数量的多少得出结论，一般不进行文献评价，也不考虑文献的质量。

### 传统文献综述的主要问题

- 传统文献评价的结果必然存在两个问题：
  - 一是多个同类研究的质量不相同；
  - 二是各个研究的样本含量的大小不相等。
- 因此，传统文献综述所采用的等权重方法很难保证研究结果真实性、可靠性和科学性，尤其当多个研究的结果不一致时，其结论容易使人产生误解或困惑。

## Meta分析的统计目的

- 对多个同类独立研究的结果进行汇总和合并分析，以达到增大样本含量，提高检验效能的目的，尤其是当多个研究结果不一致或都没有统计学意义时，采用Meta分析可得到更加接近真实情况的统计分析结果。

## 三、Meta分析的统计学过程

### ■ Meta分析统计过程的主要内容：

- 1. 异质性分析及处理
- 2. 计算合并效应量
- 3. 合并效应量的检验  
可信区间、Z检验

## (一) 异质性分析及处理

## 什么是异质性？

- 在Meta分析过程中，纳入的多个研究尽管都是对同一临床问题或具有相同研究假设的研究。但是，这些研究在纳入和排除标准、样本含量、质量控制等方面很可能不相同，从而导致了同一结局指标在多个研究间有差异。

## 异质性的种类

- Cochrane协作网将异质性分为：临床异质性、方法学异质性和统计学异质性。
- 临床异质性是指：多个研究间的受试对象不同、干预措施的差异和结局指标的变异所致的偏倚；
- 方法学异质性是指：由多个研究间的试验设计和研究质量的差异引起的异质。如盲法和分配隐藏的应用，或由于试验过程中对结局指标的定义或测量的不一致而出现的偏倚；
- 统计学异质性是指：干预效果的评价在不同试验间的变异，它是多个研究间的临床和方法学上变异联合作用的数据表现。

## 统计学异质性

- 通常，将Meta分析的统计学异质性简称为“异质性”。在Meta分析时，其异质性检验可用于判断某结局指标在多个纳入研究间是否存在异质性，而用 $I^2$ 来度量其异质性的大小。
- 各个研究的可信区间(CI)的重合程度也可以直观的反映异质性的情况，多个研究间的CI重合程度越大，存在异质性的可能性就越小，反之，其可能性就越大。
- 临床异质性、方法学异质性和统计学异质性三者是相互独立又相互关联的，临床或方法学上的异质，不一定在统计学上就有异质性，反之亦然。

## 异质性分析的意义

- Meta分析的核心计算是合并（相加），按统计原理，只有同质的资料才能进行合并或比较等统计分析，反之，则不能。

## 异质性检验

- 异质性检验  
(tests for heterogeneity)  
又称同质性检验  
(tests for homogeneity)
- 用假设检验的方法检验多个独立研究的异质性(同质性)是否具有统计学意义。

- 异质性检验的方法，目前，多由由下式计算：

$$Q = \sum W_i (d_i - \bar{d})^2 = \sum W_i d_i^2 - \frac{(\sum W_i d_i)^2}{\sum W_i}$$

$W_i$ 为每个研究的权重，第*i*个研究的权重  $W_i$ 按下式计算：

$$W_i = \frac{1}{\text{Var}(d_i)}$$

- ✓ 该检验统计量Q服从自由度为K-1的卡方 ( $\chi^2$ ) 分布，因此，当计算得到Q后，需由卡方分析获取概率，故又将此检验叫做卡方检验(Chi square test)。

- 若异质性检验检验结果为 $P > 0.10$ 时，多个研究的异质性无统计学意义；
- 若多个研究结果为 $P \leq 0.10$ 时，多个研究的异质性有统计学意义。

## $I^2$ 的计算

- 描述多个研究间的异质性大小 $I^2$ ，其计算公式如下：

$$I^2 = \frac{Q - (k - 1)}{Q} \times 100\%$$

- ✓ 式中的Q为异质性检验的卡方值 $\chi^2$ ，K为纳入Meta分析的研究个数。

## $I^2$ 的意义

- $I^2$ 可用于衡量多个研究结果间异质程度的大小。这个指标用于描述由各个研究间由非抽样误差所引起的变异（异质性）占总变异的百分比。
- 在Cochrane系统评价中，只要 $I^2$ 不大于50%，其异质性可以接受。

## 异质性与处理方法(1)

- 多个临床试验的PICOS要素不同，不能纳入同一个Meta分析。临床试验的PICOS要素：
  - 研究对象(Types of Participants): 某疾病患者
  - 干预措施(Types of Interventions): 试验组使用的某防治措施，如，药物、手术或技术等；
  - 对照措施(Types of Comparison): 对照组使用的措施，如，安慰剂、阳性对照等
  - 结局指标(Types of Outcome measures): 结局指标，如，死亡率、病死率、发病率、并发症的发生率等客观终点指标。
  - 设计方案(Types of Study): 成组，交叉等设计。

## 异质性与处理方法(2)

- 多个研究间试验设计方案(Types of Study)不同时，不能进入同一个Meta分析。如，不能将前瞻性、回顾性和横断面等研究的结果同时纳入同一个Meta分析。而应按设计类型分类，用多个Meta完成其分析。因为，不同的试验设计方案具有方法学异质性。

## 异质性与处理方法(3)

- 如果多个研究的研究对象(Types of Participants)不同，不能进入同一个Meta分析
- 如果多个研究的研究对象相同，而研究对象的基本特征不同，如性别、年龄、种族等，也需要使用亚组分析(subgroup analysis)或Meta回归(Meta-regression)进行分析

## 异质性与处理方法(4)

- 如果多个研究的干预措施(Types of Interventions)不同，不能进入同一个Meta分析
- 如果干预措施相同，而疗程长短、用药剂量、病情轻重等不同时，可使用亚组分析(subgroup analysis)或Meta回归(Meta-regression)进行分析

## 异质性与处理方法(5)

- 如果多个研究的对照措施(Types of Comparison)不同时，不能进入同一个Meta分析。如，安慰剂和阳性对照不能合并进入同一个Meta分析
- 不同的结局指标(Types of Outcome measures)，用多个Meta分析来完成

### 异质性与处理方法(6)

- 如果在Meta分析中，当异质性检验结果为  $P \leq 0.10$  时，应检查每个纳入研究的数据提取方法是否正确，数据在记录、录入过程是否准确无误。

### 异质性与处理方法(7)

- 如果在进行Meta分析中，其异质性检验出现  $P \leq 0.10$  时，还可用敏感性分析，排除可能是导致异质性的某些(个)研究后，重做Meta分析，与未排除这些(个)研究的Meta分析结果比较，探讨被去除的研究对合并效应的影响，通过比较了解其异质性的来源。

### 异质性与处理方法(8)

- 通过临床专业知识、统计学知识和前述的异质性分析方法，仍然无法解释产生异质性原因，可以采用随机效应模型(*random effect model*) 进行Meta分析
- 不用Meta分析，如果是系统评价，可做描述性(定性)的系统评价

### 不做Meta分析的情况

- 研究间的异质性无法得到合理的解释
- 多个研究的合并结果无临床意义
- 没有足够的、真实的相关研究结果
- 请记住：  
误导的结果比没有结果更糟！  
宁愿没有结果，也不要错误的结果！

## (二) 合并效应量的计算

### 多个实验效应的合并

- 将多个独立研究的结果合并成一个汇总统计量 (*overall effect*、*effect size*) 或效应尺度 (*effect magnitude*)，即用多个独立研究的某个指标的合并统计量反映其试验效应。

## 合并统计量的两种统计模型

- 固定效应模型( fixed effect model):  
若多个研究具有同质性(无异质性)时,可使用,可使用固定效应模型。
- 随机效应模型(random effect model):  
若多个研究不具有同质性时,先对异质原因进行处理,若异质性分析与处理后仍无法解决异质性时,可使用随机效应模型。

## 随机效应模型(1)

- 随机效应模型一种对异质性资料进行Meta分析的方法,但是,该法不能控制混杂、也不能校正偏倚或减少异质性,更不能消除产生异质性的原因。
- 目前,随机效应模型多采用D-L法(DerSimonian & Laird法)。

## 随机效应模型(2)

- D-L法是1986年,由DerSimonian和Laird首先提出,此法是通过增大小样本资料的权重,减少大样本资料的权重来处理资料间的异质性,而这种处理存在着较大风险。
- 通常小样本资料往往质量较差,偏倚较大,而大样本资料往往质量较好,偏倚较小。因此,经随机效应模型处理的结果,可能削弱了质量好的大样本信息,增大了质量差的小样本信息,故不应随意使用随机效应模型,其结论也应当非常谨慎和委婉。

## 分类变量

(category, dichotomous)

- 分类变量的Meta分析可选择的统计学指标:  
RD、RR和OR
- 固定效应模型(fixed effect model)
  - (1) Mantel-Haenzel法
  - (2) Peto 法
  - (3) Inverse Variance (反推方差法)
- 随机效应模型(random effects model)  
DerSimonian&Laird(D-L)法

## 数值变量

(continuous)

- 数值变量的Meta分析可选择的统计学指标:
  - (1) MD法(均数差法, Mean Difference)
  - (2) SMD法(标准化均数差法, Standardised Mean Difference)
- 固定效应模型(fixed effect model)  
Inverse Variance (反推方差法)
- 随机效应模型 (random effects model)  
DerSimonian&Laird(D-L)法

## 分类变量的统计指标

## 率差RD

- 两个发生率的差即为率差，也称危险差 (rate difference, risk difference, RD)，如，试验组发生率 (EER) 与对照组发生率 (CER) 的差，其大小可反映试验组发生率比对照组多或少绝对量。其计算公式如下：

- $$RD = EER - CER$$

## RD的意义

- 当RD=0时，可认为试验组的发生率与对照组的发生率相同
- 当RD>0时，可认为试验组的发生率大于对照组 (EER>CER)
- 当RD<0时，可认为试验组的发生率小于对照组 (EER<CER)

## 实例分析

阿斯匹林治疗心肌梗死的效果

组别	死亡	未死亡	合计
试验组	15(a)	110(b)	125(n <sub>1</sub> )
对照组	30(c)	90(d)	120(n <sub>2</sub> )
合计	45	200	245(N)

该例的 RD 为：

$$RD = EER - CER = 0.12 - 0.25 = -0.13$$

即该例试验组的病死率比对照组少 13%

无统计学意义

试验组<对照组

试验组>对照组

0

无效值

RD的可信区间的图示

## 相对危险度RR

- 相对危险度 (relative risk, RR) 也叫危险比 (risk ratio) 是前瞻性研究 (如 RCT, 队列研究等) 中较常用的指标，它是试验组 (暴露组) 某事件的发生率 EER 与对照组 (非暴露组) 某事件的发生率 CER 之比，用于说明试验组某事件的发生率是对照组的多少倍，也常用来表示暴露与疾病联系的强度及其在病因学上的意义大小。

## RR的计算

- 试验组的发生率为：EER = a/(a+b)
- 对照组的发生率为：CER = c/(c+d)
- 两个率的比值按下式计算：

$$RR = \frac{EER}{CER} = \frac{a/(a+b)}{c/(c+d)}$$

## RR的意义

- 当RR=1时, 可认为试验组的发生率与对照组的发生率相同
- 当RR>1时, 可认为试验组的发生率大于对照组(EER>CER)
- 当RR<1时, 可认为试验组的发生率小于对照组(EER<CER)

## 实例分析

阿斯匹林治疗心肌梗死的效果

组别	死亡	未死亡	合计
试验组	15(a)	110(b)	125(n <sub>1</sub> )
对照组	30(c)	90(d)	120(n <sub>2</sub> )
合计	45	200	245(N)

该例的RR为:

$$RR = \frac{EER}{CER} = \frac{a/(a+b)}{c/(c+d)} = \frac{15/125}{30/120} = 0.48$$

## 比值比OR

- 在回顾性研究(如病例对照研究)中, 往往无法得到某事件的发生率(如死亡率、病死率、发病率), 也就无法计算出RR。但是当该发生率很低时(如发生率小于或等于5%), 可以计算出一个RR的近似值, 该近似值称为OR, 即是比值比(odds ratio)。

## OR的计算(1)

- 病例组的暴露与未暴露的比值:  
 $odds_1 = a/b$
- 对照组的暴露与未暴露的比值  
 $odds_0 = c/d$

## OR的计算(2)

- 试验组的比值(odds<sub>1</sub>)与对照组(odds<sub>0</sub>)的比值之比即为比值比(odds ratio、比数比、机会比、优势比), 其计算公式为:

$$OR = \frac{odds_1}{odds_0} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

## OR的意义

- 当OR=1时, 可认为病例组的比值与对照组的比值相同
- 当OR>1时, 可认为病例组的比值大于对照组(odds<sub>1</sub>>odds<sub>0</sub>)
- 当OR<1时, 可认为病例组的比值小于对照组(odds<sub>1</sub><odds<sub>0</sub>)

## RR与OR的选择

- 在前瞻性研究中(如RCT), 可计算RR, 若某事件的发生率较低时(如发生率小于5%), 可用OR估计RR。如果某事件的发生率较大时, OR与RR相差较大。
- 而在回顾性研究(如, 病例对照研究)中, 由于无法计算发生率、死亡率等指标, 也就无法计算RR, 只能用OR。

## 数值变量的统计指标

## 数值变量的指标-MD

- 均数差(加权均数差)
- (Mean Difference, MD or Weighted Mean Difference WMD)
- 某个研究的MD可按式计算:

$$MD = \bar{X}_1 - \bar{X}_2$$

$\bar{X}_1$ 为试验组的均数,  $\bar{X}_2$ 为对照组的均数

- 从公式可见, 均数差(MD)即为两均数的差值。
- 该指标以试验原有的测量单位, 真实地反映了试验效应, 消除了绝对值大小对结果的影响, 在实际应用时, 该指标容易被理解和解释。

## 数值变量的指标-SMD

- 标准化均数差
- (Standardised Mean Difference, SMD)
- 某个研究的标准化均数差SMD, 可按式计算:

$$SMD = \frac{\bar{X}_1 - \bar{X}_2}{S_C} \left(1 - \frac{3}{4N-9}\right)$$

$$S_C = \sqrt{\frac{S_1^2(n_1-1) + S_2^2(n_2-1)}{(n_1+n_2-2)}}$$

- SMD可简单地理解为两均数的差值再除以合并标准差的商, 它不仅消除了某研究的绝对值大小的影响, 还消除了测量单位对结果的影响。因此, 该指标尤其适用于单位不同或均数相差较大的数值资料分析。
- 但是, 标准化均数差(SMD)是一个没有单位的值, 因而, 对SMD分析的结果解释要慎重。

### (三) 合并效应量的检验

### 合并效应量的检验

- 用假设检验(hypothesis test)的方法检验多个独立研究的总效应量(效应尺度)是否具有统计学意义,其原理与常规的假设检验完全相同。
- 两种方法:
  - ①u检验(Z test)
  - ②卡方检验(Chi square test)

- 根据z或(u)值或卡方值得到该统计量下概率(P)值。
- 若 $P \leq 0.05$ ,多个研究的合并效应量有统计学意义;
- 若 $P > 0.05$ ,多个研究的合并效应量没有统计学意义。

### 合并效应量的可信区间

- 在Meta分析中,常用可信区间进行假设检验,95%的可信区间与 $\alpha$ 为0.05的假设检验等价,99%的可信区间与 $\alpha$ 为0.01的假设检验等价。
- 此外,森林图即是根据各个独立研究的95%可信区间及合并效应量的95%可信区间绘制的。

### 合并效应量的可信区间

- 合并统计量的95%可信区间下限:
$$\theta - 1.96 \times SE(\theta)$$
- 合并统计量的95%可信区间上限:
$$\theta + 1.96 \times SE(\theta)$$

式中的 $\theta$ 为某合并统计量,如 $MD_{\text{合并}}$ 、 $\ln(RR_{\text{合并}})$

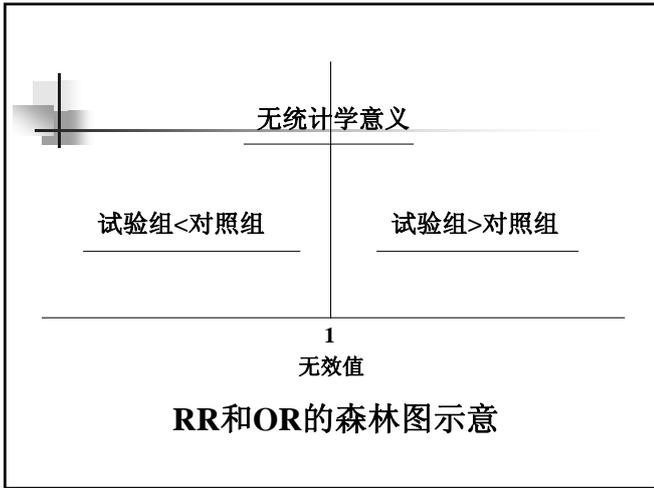
试验组<对照组

试验组>对照组

无统计学意义

0  
无效值

RD、MD与SMD的森林图示意



## 四、分类变量的实例分析

### 单个分类变量的研究数据

- 分类变量(category, dichotomous)单个研究的统计量 $d_i$ ，可选择OR、RR或RD，四格表数据如下表

	Event	No Event	
Treatment	$a_i$	$b_i$	$n_{1i}$
Control	$c_i$	$d_i$	$n_{2i}$
	$m_{1i}$	$m_{2i}$	$N_i$

### K研究的分类变量资料整理

i 个 研究	Treatment			Control		
	Event	No Event		Event	No Event	
1	$a_1$	$b_1$	$n_{11}$	$c_1$	$d_1$	$n_{21}$
2	$a_2$	$b_2$	$n_{12}$	$c_2$	$d_2$	$n_{22}$
3	$a_3$	$b_3$	$n_{13}$	$c_3$	$d_3$	$n_{23}$
...	...	...	...	...	...	...

### 实例一

七个阿司匹林预防心肌梗死的研究资料

(Fleiss JL, Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. J Clin Epidemiol. 1991;44(2):127-39).

K 个 研究	阿司匹林		安慰剂		合计 $N_i$	OR	95% CI	
	死亡数 $a_i$	未死亡数 $b_i$	死亡数 $c_i$	未死亡数 $d_i$			下限	上限
1	49	566	67	557	1239	0.720	0.489	1.059
2	44	714	64	707	1529	0.681	0.457	1.013
3	102	730	126	724	1682	0.803	0.606	1.063
4	32	285	38	271	626	0.801	0.486	1.319
5	85	725	52	354	1216	0.798	0.553	1.153
6	246	2021	219	2038	4524	1.133	0.935	1.373
7	1570	7017	1720	6880	17187	0.895	0.829	0.966
合计	2128	12058	2286	11531	28003			

表中 $a_i$ 、 $b_i$ 、 $c_i$ 、 $d_i$ 为各研究四格表数， $N_i$ 为各个研究的样本例数。

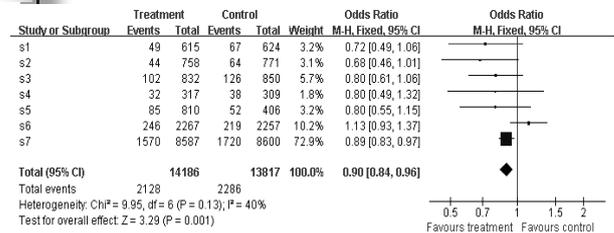
### 二分类资料的数据录入界面

Study or Subgroup	Treatment Events	Total	Control Events	Total	Weight	Odds Ratio M-H, Fixed, 95% CI
1	49	615	67	624	3.2%	0.72 [0.49, 1.06]
2	44	750	64	771	3.1%	0.68 [0.46, 1.01]
3	102	837	126	950	5.7%	0.80 [0.61, 1.06]
4	22	317	38	309	1.9%	0.80 [0.49, 1.32]
5	85	810	52	406	3.2%	0.80 [0.55, 1.15]
6	246	2267	219	2257	10.2%	1.13 [0.93, 1.37]
7	1570	8987	1720	8600	72.9%	0.89 [0.83, 0.97]
Total (95% CI)		14186		13817	100.0%	0.80 [0.84, 0.96]
Total events		2128		2286		
Heterogeneity: Chi <sup>2</sup> = 9.95, df = 6 (P = 0.12); I <sup>2</sup> = 40%						
Test for overall effect: Z = 3.29 (P = 0.001)						

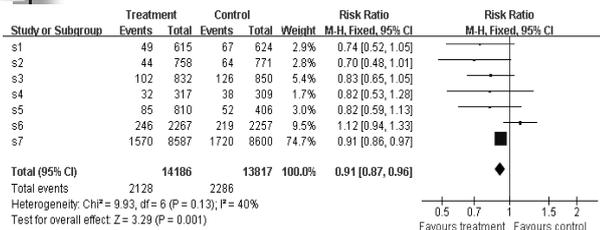
## RR和OR的森林图

- RR和OR的森林图(forest plots), 无效线竖线的横轴尺度为1, 每条横线为该研究的95%可信区间上下限的连线, 其线条长短直观地表示了可信区间范围的大小, 线条中央的小方块为RR或OR值的位置, 其方块大小为该研究权重大小。若某个研究95%可信区间的线条横跨为无效竖线, 即该研究无统计学意义, 反之, 若该横线落在无效竖线的左侧或右侧, 该研究有统计学意义。

## 例一的Revman5.2 森林图



## 例一的Revman5.2 森林图



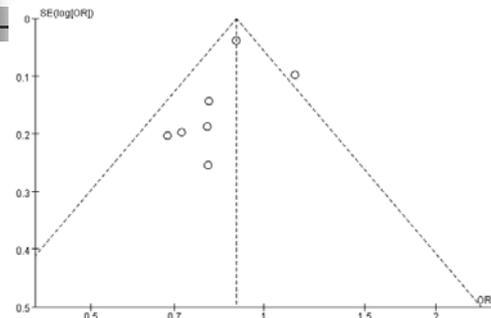
## 漏斗图及用途

- 漏斗图 (funnel plots) 最初是用每个研究的处理效应估计值为X轴, 样本含量的大小为Y轴的简单散点图 (scatter plots)。
- 对处理效应的估计, 其准确性是随样本含量的增加而增加, 小样本研究的效应估计值分布于图的底部, 其分布范围较宽; 大样本研究的效应估计值分布范围较窄, 当没有发表偏倚时, 其图形呈对称的倒漏斗状, 故称之为“漏斗图”。

## RevMan 中的漏斗图

- 在RevMan软件中, 漏斗图是采用RR或OR对数值 (logOR或logRR) 为横坐标, OR或RR对数值标准误的倒数1/SE(logRR)为纵坐标绘制的, 然后, 以真数标明横坐标的标尺, 而以SE(logRR)标明纵坐标的标尺。

## 例一的Revman5.2 漏斗图 (Funnel Plot)



## 漏斗图的用途

- 漏斗图主要用于观察某个系统评价或Meta分析结果是否存在偏倚，如发表偏倚或其他偏倚。如果资料存在偏倚，会出现不对称的漏斗图，不对称越明显，偏倚程度也就越大。漏斗图的不对称性主要与发表偏倚有关，但也可能存在其他原因。

## 漏斗图不对称的主要原因

- 导致漏斗图不对称的主要原因有：
  - 发表偏倚 (Publication bias)
  - 选择性偏倚 (Selection bias)
  - 语言偏倚 (Language bias)
  - 引用偏倚 (Citation bias)
  - 重复发表偏倚 (Multiple publication bias)
  - 等等

## 五、数值变量的实例分析

### 单个数值变量研究的数据

- 数值变量 (continuous) 单个研究的统计量  $d_i$ ，可选择WMD和SMD法，单个研究的数据如下表

	例数	均数	标准差
	$n_i$	$\bar{X}_i$	$S_i$
试验组	$n_1$	$\bar{X}_1$	$S_1$
对照组	$n_2$	$\bar{X}_2$	$S_2$

### K研究的数值变量资料整理

第i个研究	试验组			对照组			$N_i$
	均数	标准差	例数	均数	标准差	例数	
	$\bar{X}_{1i}$	$S_{1i}$	$n_{1i}$	$\bar{X}_{2i}$	$S_{2i}$	$n_{2i}$	
$i=1$	$\bar{X}_{11}$	$S_{11}$	$n_{11}$	$\bar{X}_{21}$	$S_{21}$	$n_{21}$	$N_1$
$i=2$	$\bar{X}_{12}$	$S_{12}$	$n_{12}$	$\bar{X}_{22}$	$S_{22}$	$n_{22}$	$N_2$
$i=3$	$\bar{X}_{13}$	$S_{13}$	$n_{13}$	$\bar{X}_{23}$	$S_{23}$	$n_{23}$	$N_3$
...	...	...	...	...	...	...	...

### 实例二

女童掌骨II皮质厚度的11个研究

研究	K个	高氟区			低氟区			P值
		$n_1$	$\bar{X}_1$	$S_1$	$n_2$	$\bar{X}_2$	$S_2$	
1	26	2.26	0.32	42	2.33	0.33	P>0.05	
2	55	2.39	0.31	40	2.49	0.32	P>0.05	
3	46	2.50	0.30	50	2.67	0.35	P<0.05	
4	45	2.64	0.26	50	2.90	0.45	P<0.05	
5	45	2.81	0.35	45	2.93	0.36	P>0.05	
6	52	2.95	0.46	55	3.27	0.37	P<0.05	
7	46	3.15	0.39	42	3.48	0.48	P<0.05	
8	45	3.47	0.46	51	3.73	0.54	P<0.05	
9	45	3.63	0.38	45	3.81	0.40	P<0.05	
10	42	3.81	0.41	45	4.16	0.42	P<0.05	
11	44	3.99	0.56	25	4.18	0.41	P>0.05	
合计	491			490				

方积乾.医学统计学与电脑实验.第二版.上海科学技术出版社.2001.349-350.

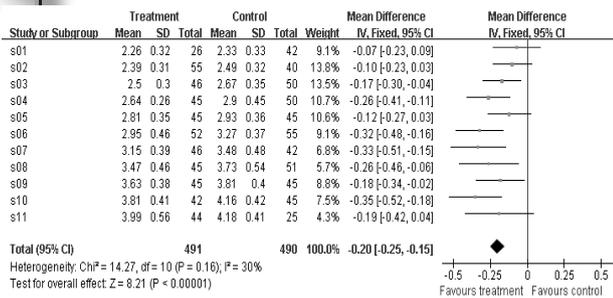
## 数值资料的数据录入界面

Study or Subgroup	Treatment			Control			Weight	Mean Difference IV, Fixed, 95% CI
	Mean	SD	Total	Mean	SD	Total		
s01	2.26	0.32	26	2.33	0.33	42	9.1%	-0.07 [-0.23, 0.09]
s02	2.39	0.31	55	2.49	0.32	40	13.6%	-0.10 [-0.23, 0.03]
s03	2.5	0.3	46	2.67	0.35	50	13.5%	-0.17 [-0.30, -0.04]
s04	2.64	0.26	45	2.9	0.45	50	10.7%	-0.26 [-0.41, -0.11]
s05	2.81	0.35	45	2.93	0.36	45	10.6%	-0.12 [-0.27, 0.03]
s06	2.95	0.46	52	3.27	0.37	55	9.1%	-0.32 [-0.48, -0.16]
s07	3.15	0.39	46	3.48	0.48	42	6.8%	-0.33 [-0.51, -0.15]
s08	3.47	0.46	45	3.73	0.54	51	5.7%	-0.26 [-0.46, -0.06]
s09	3.63	0.38	45	3.81	0.4	45	8.8%	-0.18 [-0.34, -0.02]
s10	3.81	0.41	42	4.16	0.42	45	7.5%	-0.35 [-0.52, -0.18]
s11	3.99	0.56	44	4.18	0.41	25	4.3%	-0.19 [-0.42, 0.04]
Total (95% CI)			491			490	100.0%	-0.20 [-0.25, -0.15]
Heterogeneity: Chi <sup>2</sup> = 14.27, df = 10 (P = 0.16); I <sup>2</sup> = 30%								
Test for overall effect: Z = 8.21 (P < 0.00001)								

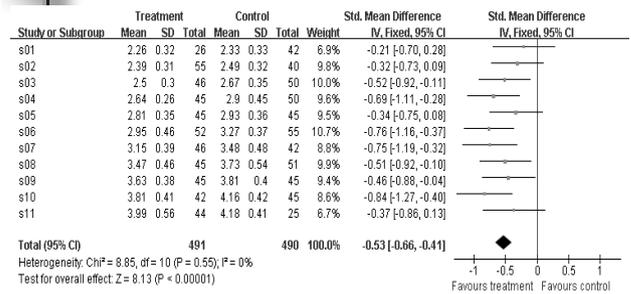
## MD和SMD的森林图

- MD和SMD的森林图，无效线竖线的横轴尺度为0，每条横线为该研究的95%可信区间上下限的连线，其线条长短直观地表示了可信区间范围的大小，线条中央的小方块为MD或SMD值的位置，其方块大小为该研究权重大小。若某个研究95%可信区间的线条横跨为无效竖线，即该研究无统计学意义，反之，若该横线落在无效竖线的左侧或右侧，该研究有统计学意义。

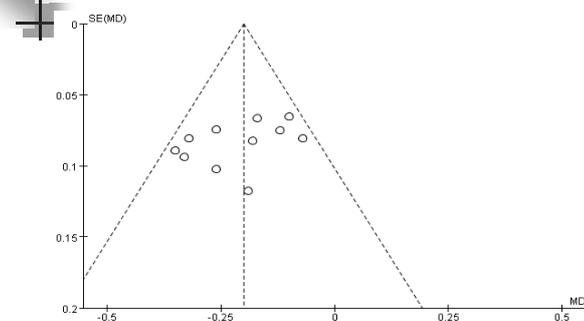
## 例二的Revman5.2森林图 (MD法)



## 例二的Revman 5.2森林图 (SMD法)



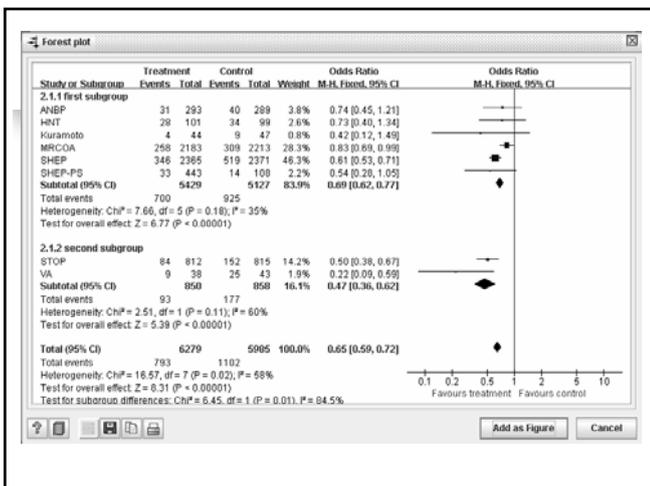
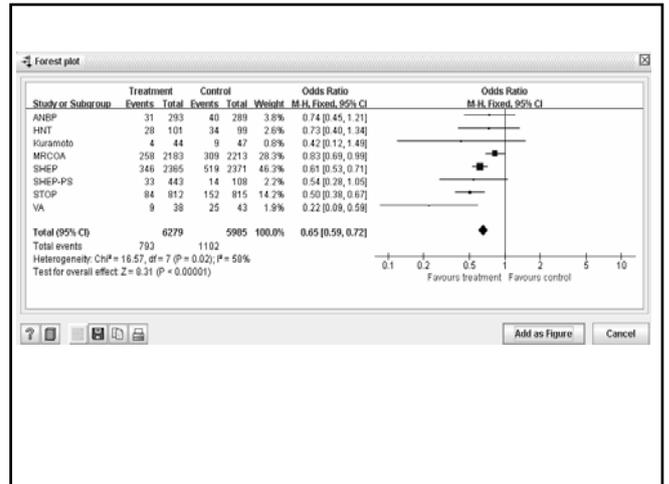
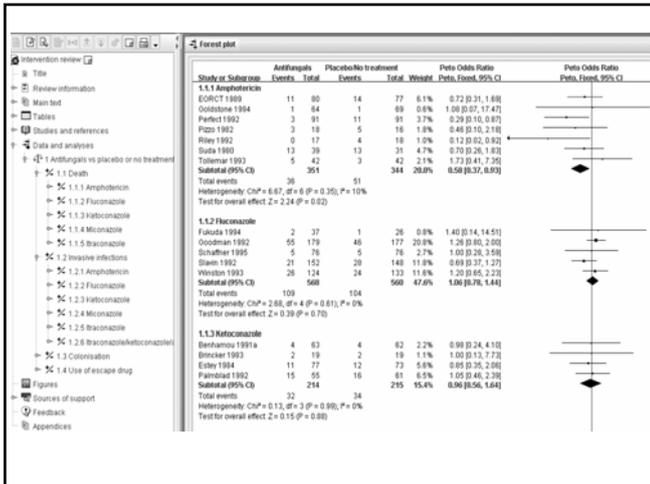
## 例二的Revman5.2漏斗图 (Funnel Plot)



## 六、亚组分析

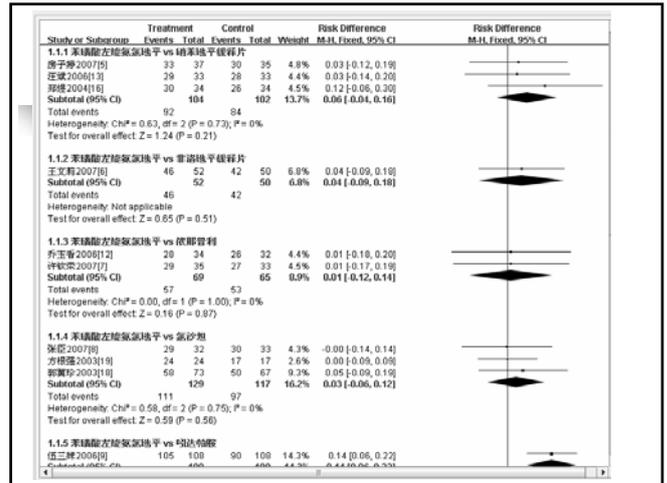
- 亚组分析(subgroup analysis), 即分层分析。在Meta分析是按某个(些)特征或属性将纳入的多个研究分为若干个小单元进行分类合并, 如前述的不同的干预措施, 不同的对照都需要做亚组分析, 以消除这些不同(异质)对试验效应组间比较的影响。

- 例一:
  - 多个研究的干预措施不同, 只要有临床异质性, 无论其异质性检验有无统计学意义, 都应使用亚组分析(subgroup analysis), 见分析结果。



- 多个研究的干预措施不同, 如果没有临床异质性, 其异质性检验有统计学意义, 最好是对异质性进行分析, 找到异质性来源后, 再决定是否使用亚组分析。

- 例二：
- 多个研究的对照措施不同，只要有临床异质性，无论其异质性检验有无统计学意义，也应使用亚组分析 (subgroup analysis)，见分析结果。



- 多个研究的对照措施不同，如果没有临床异质性，其异质性检验有统计学意义，最好也是对异质性进行分析，找到异质性来源后，再决定是否使用亚组分析。

## 七、有关Meta分析的讨论

### 关于Meta分析的争论

- 对Meta分析的争论也较多，主要如下
- 肯定Meta分析者间的争论：
  - 1.对固定与随机效应模型的争论；
  - 2.权重(W)计算的不同方法等。
- 否定Meta分析者

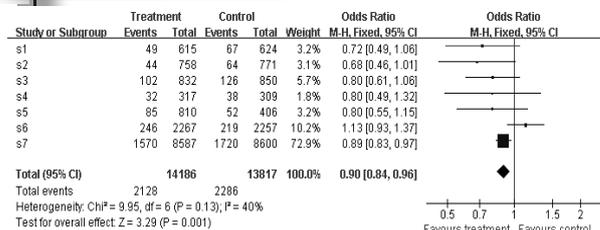
### 关于随机效应模型 (1)

- 随机效应模型最主要用于：
  - 1、满足固定效应模型Meta分析的条件时，采用随机效应模型报道其结果，可使分析结论更为谨慎、委婉和可靠。
  - 2、根据异质性分析，确认其异质性主要来源于方法学偏倚，且合并效应量对临床或对未来研究有较好的参考意义时，可使用随机效应模型进行探索性分析。

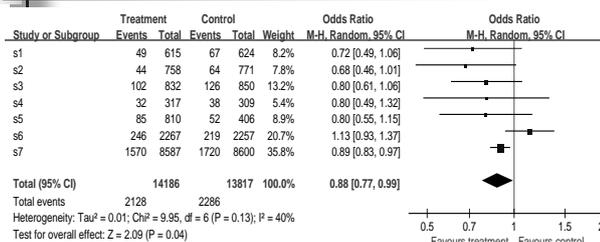
## 关于随机效应模型(2)

- 以下情况不要使用随机效应模型：
  - 1、多个研究间的PICOS不同，或不满足同质性条件时，不能使用随机效应模型。
  - 2、未做研究间的异质性分析，或研究间的异质性来源不明，或异质性无法得到合理的解释时，不宜使用随机效应模型。
  - 3、多个研究的合并结果无临床意义。

## 例一的固定效应模型



## 例一的随机效应模型



## Meta分析的局限性

- 目前，Meta分析的统计方法尚不够完善，还不能满足不同资料类型和不同的临床设计方案的需要，如多个均数比较、等级资料比较时，仍无成熟的Meta分析方法。

## 关于Meta分析的思考

- 正确应用Meta分析：
 

既不能扩大Meta分析的作用，也不能否定Meta分析的用途。
- 正确解释Meta分析的结果：
 

对任何统计分析的结果，都需要结合医学专业知识和统计学知识，对研究结果做出尽可能客观和真实的解释，Meta分析也是如此。

## 八、Meta分析的软件

- **1. Review Manager(Revman):** 该软件是国际Cochrane协作网系统评价的标准化专用软件, 其中包含了Cochrane系统评价的各项功能, 也包括该组织推荐的各种Meta分析功能, 具有操作简单、结果直观的特点, 目前最新版本为5.3.5。该软件可以从网上免费下载, 其下载地址为:
  - <http://tech.cochrane.org/revman/download>

- **2. STATA**该软件是美国Stata Corp LP研制的统计软件, 它是Meta分析功能很强大的统计软件, 从1984年起, 连续推出了多个版本, 最新版本为14.0。它是目前Meta分析功能最强的软件。(http://www.stata.com/)
- 该软件可完成二分类变量和连续性变量的Meta分析, 也可以进行Meta回归分析, 还可以绘制Meta分析的相关图型, 如森林图(Forest)、漏斗图(Funnel)和L'Abbe图

### 主要参考文献

- **1. Julian PT Higgins, Sally Green (editors).** Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011].
- **2. The Australasian Cochrane Center.** The Cochrane Handbook Summary Points & Libraries of Examples, 1998.
- **3. 李幼平主编.**循证医学(第3版).高等教育出版社, 2013.12

# 谢谢!



Tel: 18980601800  
806125243@qq.com