
Meta分析的10个问题：从理论概念到R语言操作实践

作者：谷鸿秋 来源：“统技思维”公众号

本文原地址：<https://www.iikx.com/news/article/6111.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

Meta分析的10个问题

：从理论概念到R语言操作实践。本文想对Meta分析做一些提炼和总结，为方便叙述，还是以问答形式做一个分享吧。

拟回答的问题：

Meta分析=系统综述=循证医学？

只有RCT才能做Meta分析吗？

Meta分析常用的效应指标有哪些？

Meta分析常用的软件有哪些？

Meta分析的步骤是什么？

Meta分析的结果如何解读？

单组率的meta分析R编程

连续变量效应指标的Meta分析R编程

分类变量效应指标的Meta分析R编程

Meta回归的R编程

首先，先澄清一些重要的，概念上的误解。

1.Meta分析=系统综述=循证医学？

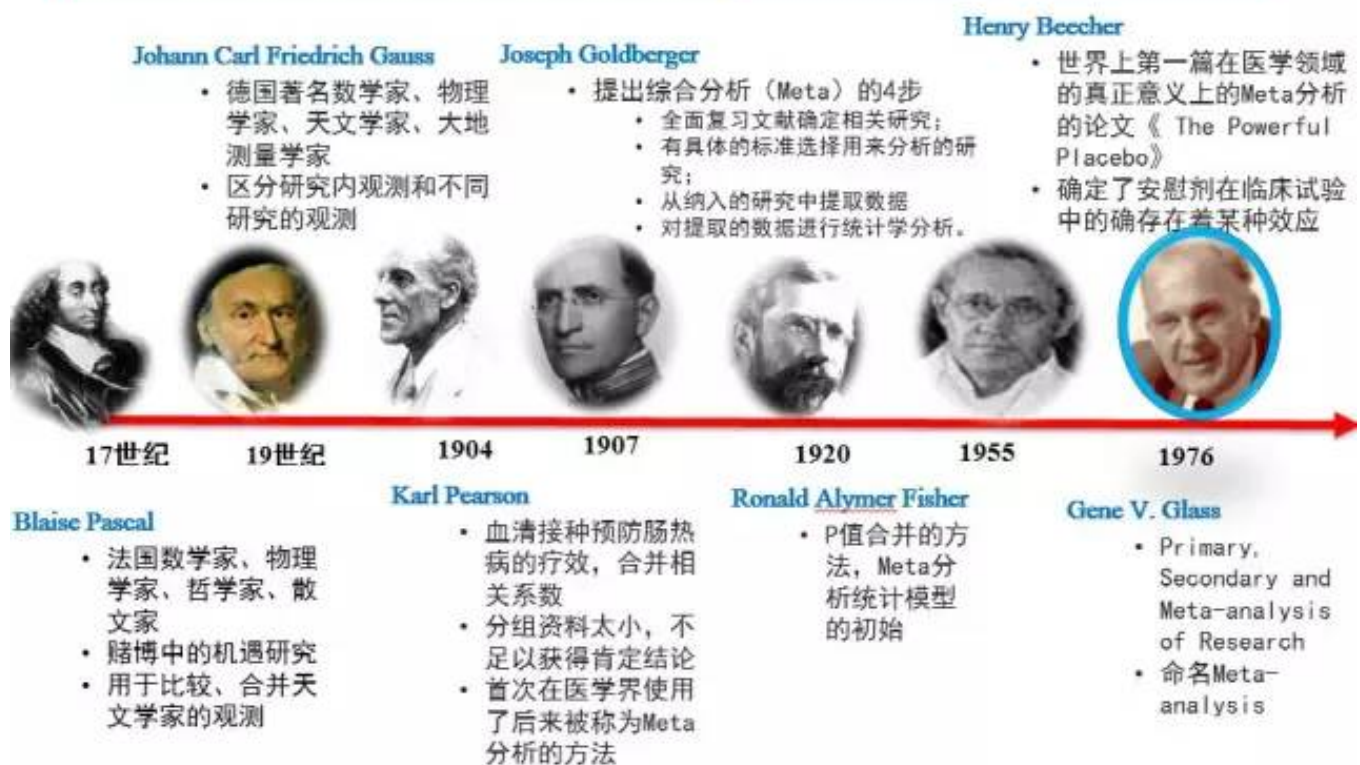
为了要说清这个问题，请不要猴急，我们先一起理一理他们各自的来龙去脉。

Meta分析(Meta-analysis, MA)，其思想最早可追溯到17世纪法国数学家Blaise Pascal对「机遇」研究，后被用于天文观测数据的比较、合并。1904年英国数学家及统计学家Karl Pearson在「血清接种预防肠热病的疗效」研究中，首次在医学界使用了后来被称为Meta分析的方法。

1907年，美国的内科医生，流行学家Joseph Goldberger提出了综合分析的4步法，勾勒出了现在的Meta分析步骤。英国生物统计学家，大家耳熟能详的Ronald Fisher在1920年提出了合并P值的方法，奠定了Meta分析统计模型的基础。不过Meta分析这一术语首次正式被提出，还是在Gene V. Glass的《Primary, Secondary and Meta-analysis of Research》这一著作中。

关于Meta分析的这段历史，可见下图：

Meta分析的缘起及发展历程



那么meta分析的定义，具体是：

Miquel Porta在流行病学词典里给出了一个较为宽泛的定义：

A statistical analysis of results from separate studies, examining sources of differences in results among studies, and leading to a quantitative summary of the results if the results are judged sufficiently similar to support such synthesis. (Miquel Porta, 2008)

而考科兰给出了一个更为严格的定义：

The use of statistical techniques in a systematic review to integrate the results of included studies. Sometimes misused as a synonym for systematic reviews, where the review includes a meta-analysis. (The Cochrane Collaboration)

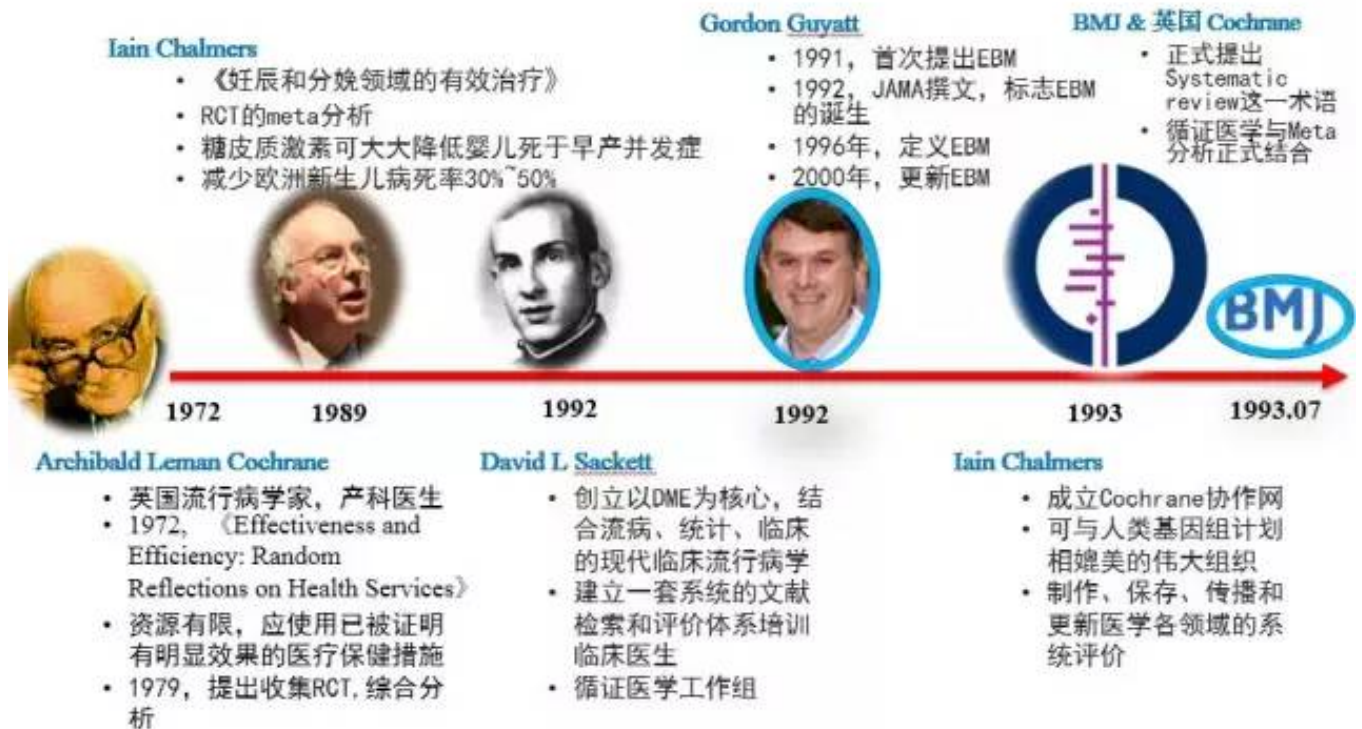
两种定义确有差别，但从本质上讲，Meta分析其实就是一种数据合并的统计方法。如果严格限定其使用场景，如只用于系统综述中，那就是考科兰的定义，如果不限定其使用场景，那就是Miguel Porta的定义。

前面提到「系统综述」，要说起清楚系统综述，还得拉上另一大家耳熟能详的术语：「循证医学」。

循证医学(Evidence-Based Medicine, EBM)，其思想源于1979年英国流行病学家，产科医生Archibald Lemmon Cochrane提出的「收集RCT,进行综合分析」这一理念，不过「循证医学」医学这一术语也是到了1991年才由加拿大的学者Gordon Guyatt首次提出，而「系统综述」则是在1993年，由BMJ期刊的编辑和英国Cochrane的专家商讨后正式提出，也就是在此时，循证医学，系统综述与Meta分析才完成了彼此的结合。

一个过程，请见下图：

循证医学的缘起及发展历程



具体说到循证医学的定义，最为经典的当属David L. Sackett给出的包含现能获得的最佳证据、临床大夫的诊疗经验以及病人价值观和期许三方面要素的定义。

The conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.

The practice of evidence-based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research.

循证医学是慎重、准确和明智地应用当前所能获得最好的研究证据，整合最佳临床证据、临床经验以及患者价值观来确定患者的治疗措施的一门学科。



那么如何获得当前最好的研究证据呢?这就是系统综述所要回答的问题。

具体到系统综述，不同的学者也给出了不同的定义，但究其含义，无非是指针对某特定的研究问题，系统深入地查找、选择、合并高质量研究证据，以寻求问题的答案。在此过程中，Meta分析可用(定量合并)，可不用(定性合并)

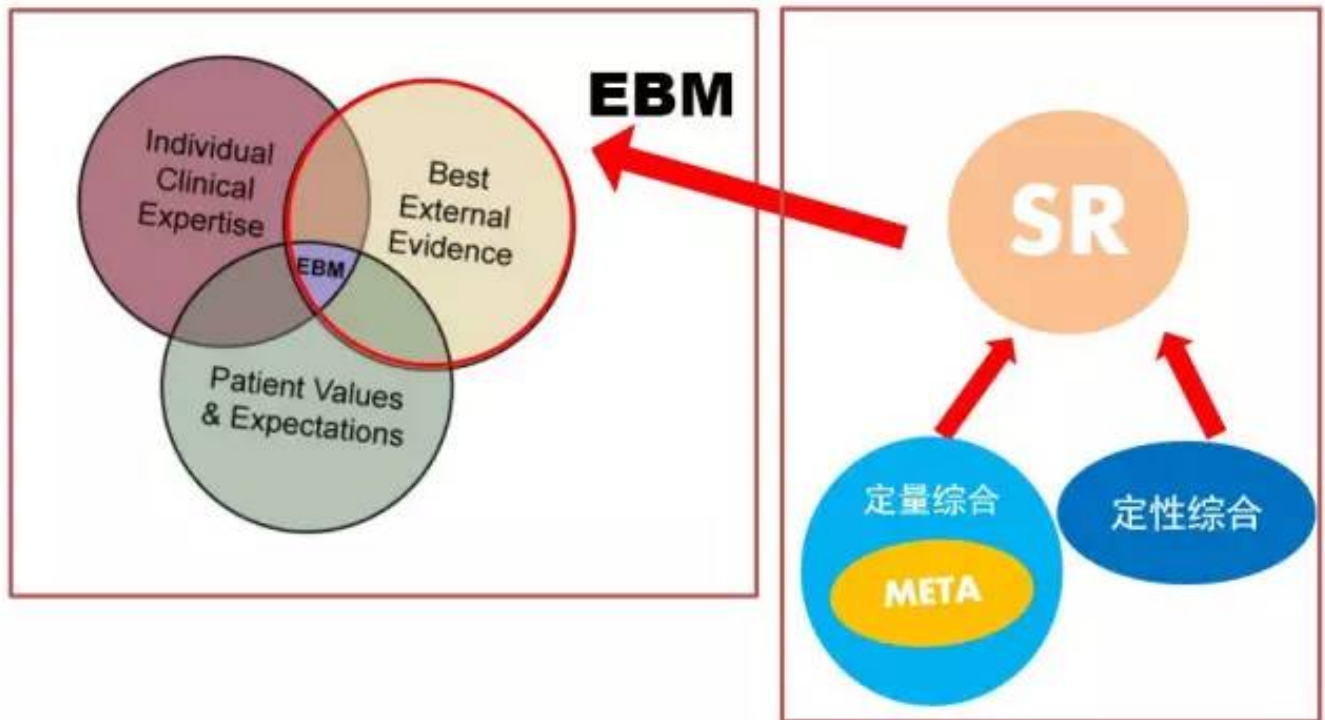
A summary of the medical literature that use explicit methods to perform a thorough literature search and critical appraisal of individual studies and that uses appropriate statistical techniques to combine these valid studies. (David L. Sackett, 2000)

The application of strategies that limit bias in the assembly, critical appraisal, and synthesis of all relevant studies on a specific topic. meta-analysis may be, but is not necessarily, used as part of this process. (Miquel Porta, 2008)

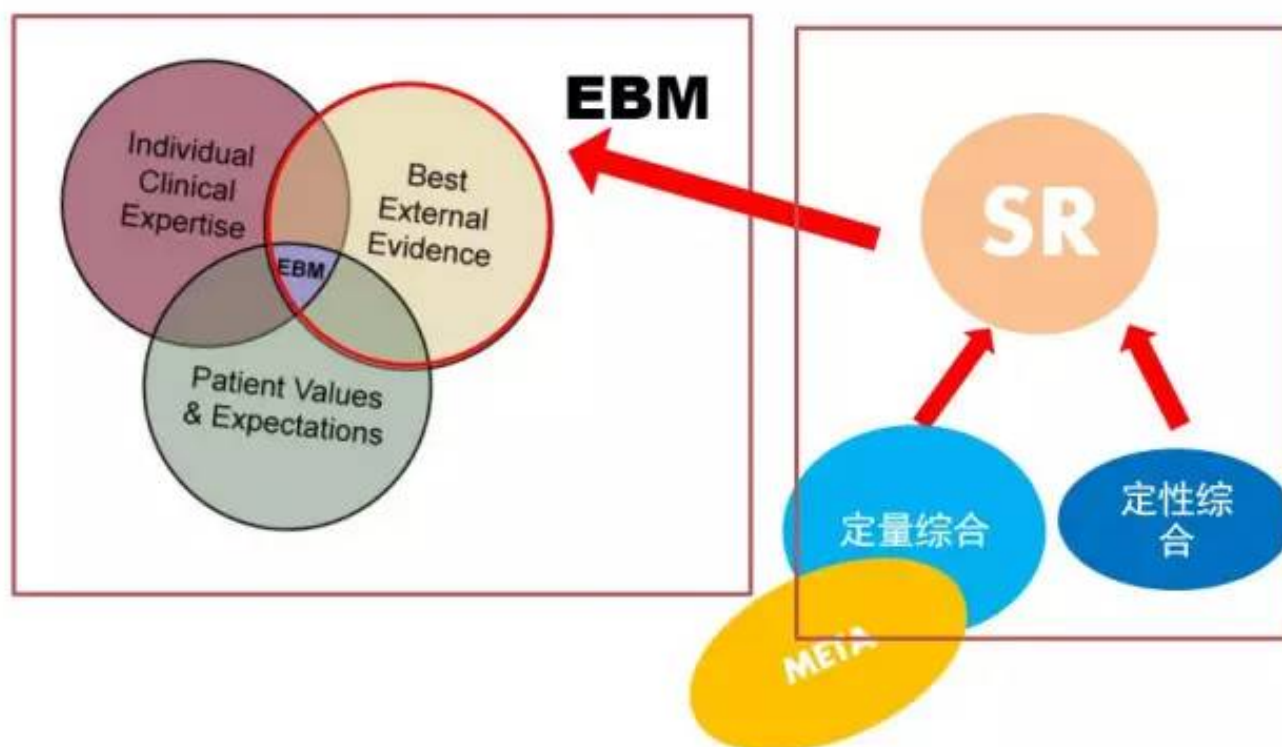
A systematic review is a high-level overview of primary research on a particular research question that tries to identify, select, synthesize and appraise all high quality research evidence relevant to that question in order to answer it. (The Cochrane Collaboration)

因此，总结而言，Meta分析(MA)、系统综述(SR)以及循证医学(EBM)的关系如下：

EBM、SR、MA-狭义关系



EBM、SR、MA-广义关系



2. 只有RCT才能做Meta分析吗?

循证医学思想的初衷是收集合并RCT研究，但随着其不但扩张延伸，非随机干预研究、观察性研究、诊断试验等都可以进行Meta分析。

Meta分析种类



3.Meta分析常用的效应指标有哪些?

依据不同的结局数据的属性，效应指标也有所不同。具体而言，归纳如下：

Meta分析相关概念

Meta分析的常用效应指标

- 分类性结局指标
 - 比值比 (Odds Ratio, OR)
 - 相对危险度 (Risk Ratio, RR)
 - 风险差 (Risk Difference, RD)
- 连续性结局指标
 - 加权均数差 (Weighted Mean Difference, WMD)
 - 变化均数差 (Standardized Mean Difference, SMD)
- 生存数据结局指标
 - 风险比 (Hazard Ratio, HR)
- 诊断试验
 - 灵敏度 (Sensitivity, Sen)
 - 特异度 (Specificity, Spe)
 - 诊断优势比 (Diagnostic odds ration, DOR)

此外，其他效应指标如：EER/CER、RRR、RRI、RBI、ARR、ARI、NNT、NNH

部分效应指标的特性归纳如下：

表 2 OR、RR和RD三者特征的比较
Table 2 Comparison of some characteristics among OR, RR and RD

	OR	RR	RD
无效值/ Null value	1	1	0
可解释性/ Interpretation	差/ Difficult	好/ Easy	好/ Easy
适合的研究类型 Study designs available	所有有对照的设计 All designs with control group	前瞻性研究 Prospective study	前瞻性研究 Prospective study
控制协变量时的应用性/ Covariate adjustment	好/ Easy	较差/ Hard	差/ Very hard
一致性/ Consistency	好/ Good	好/ Good	差/ Inconsistent
以事件或非事件计算结果时 Labeling of events and nonevents	不变 Invariant	可能差异较大 Creates ambiguity	不变 Invariant

效应指标的选择参考：

表 3 Meta分析合并统计量的选择
Table 3 Selection of summary statistics for Meta-analysis

		OR	RR	RD	WMD	SMD
流行病学设计类型 Epidemiological design	随机对照试验/ Randomised controlled trial	+	++	+	++	++
	队列研究/ Cohort study	+	++	+	++	++
	病例对照研究/ Case-control study	++	—	—	+	+
	横断面研究/ Cross-sectional study	+	—	—	+	+
资料类型 Data type	二分类变量/ Dichotomous variable	+	+	+	—	—
	连续性变量/ Continuous variable	—	—	—	+	+

*注释: ++最适合; +适合; —不恰当 Notes: ++Most appropriate; +Appropriate; —Inappropriate

4. Meta分析常用的软件有哪些?

通用的统计软件Stata, R均可以, 专用软件Revman, 以及CMA等其他专用小软件。当然, SAS其实也是可以的, 不过此处不做推荐。

Meta分析软件

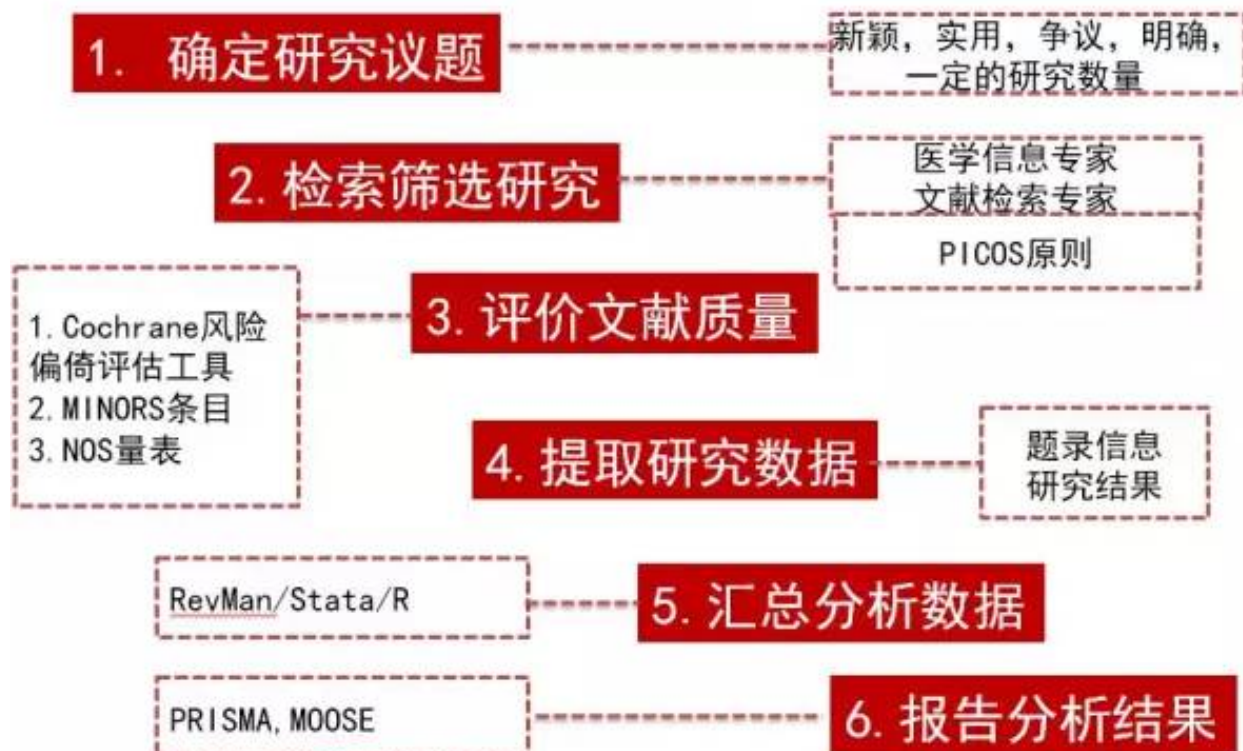
- RevMan
 - 免费
 - Win/OS/Linux
 - Cochrane官方软件
 - 图形操作界面
 - 能完成经典的Meta分析
- Stata/R
 - Stata收费, R免费
 - Win/OS/Linux
 - 编程语言操作
 - 功能强大
- 其他专用软件
 - openMA
 - CMA
 - MA
 - Meta-DiSc



5. Meta分析的步骤是什么?

这一部分, 网上资料较多, 操作讲起来也费劲, 就晒一张鸡汤式的幻灯片吧, 6步搞定Meta分析!

Meta分析上手



6. Meta分析的结果如何解读?

Meta分析的结果主要靠图形展示, 这些图中的男一号就是森林图, 女一号就是漏斗图。此外, 还有其他各种不同用途的图形。

Meta分析结果解读

Meta分析涉及的主要图形

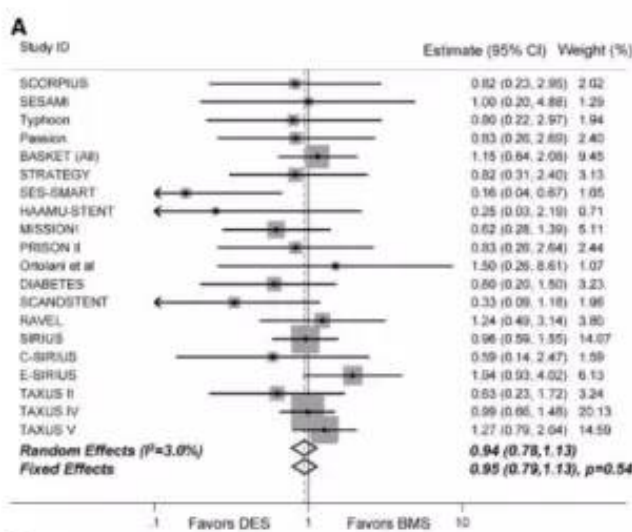
- 森林图
- 漏斗图
- 风险偏倚图
- 星状图
- 拉贝图
- 加尔布雷斯图
- Egger线性回归图
- Meta回归图

先看看男一号森林图及其解读：

森林图

森林图：以无效线（横坐标刻度为0或1）为中心，结合了数字、文本、图形，同时展示各研究以及汇总研究结果的综合图形。

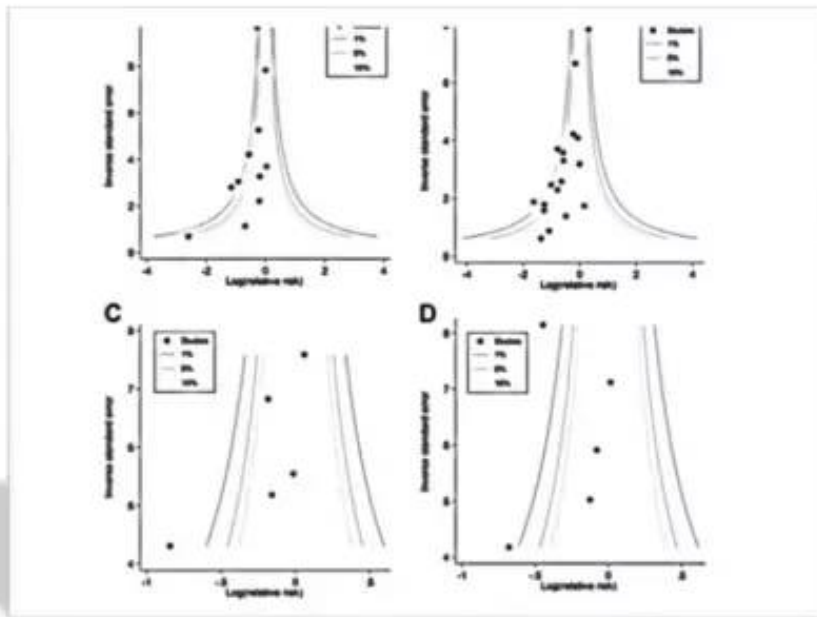
- 方框：权重
- 黑点：效应值的点估计
- 线条：效应值的置信限
- 箭头：超出显示范围
- 菱形：汇总结果
 - 中心：汇总结果点估计
 - 宽度：汇总结果置信限
- 随机效应 vs 固定效应
 - 随机效应：效应真值不同
 - 固定效应：效应真值相同



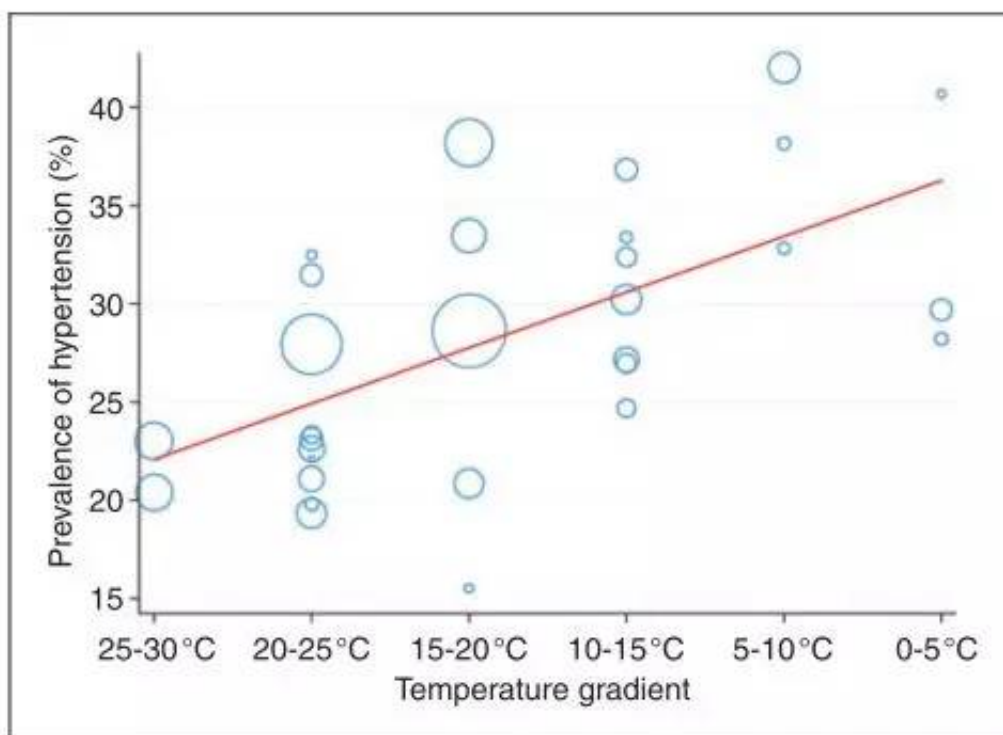
再看看女一号漏斗图:

漏斗图

漏斗图：以每项研究估计所得的效果为X轴，每项研究的样本量。大小或者精度为Y轴做出的散点图。目测是否对称以检查是否存在发表偏倚。



其他的图，在此不做具体介绍了。再看一个Meta回归图吧，Meta回归可以理解为研究层面的回归分析。



接下来是操作部分，考虑到费用、易获得性和易用性，推荐的软件首选R软件。安装R后，首选安装meta、metafor包，并加载。

```
install.packages("meta")
```

```
install.packages("metafor")
```

```
library(meta)
```

```
library(metafor)
```

7.单组率meta分析的R编程

对于一些患病率率，耐药率的研究，对于诊断试验的灵敏度、特异度的汇总，都可以考虑采用单组数据的Meta分析。

具体到R编程，可采用Meta包的metaprop函数。

```
metaprop(event, n, data=, sm=)
```

metaprop最重要的几个参数：

event:事件数;

n:总数;

data:数据来源

sm:合并率时的采用的效应指标，其实为各种率的转换方法。

一个单组数据meta分析的R编程示例：

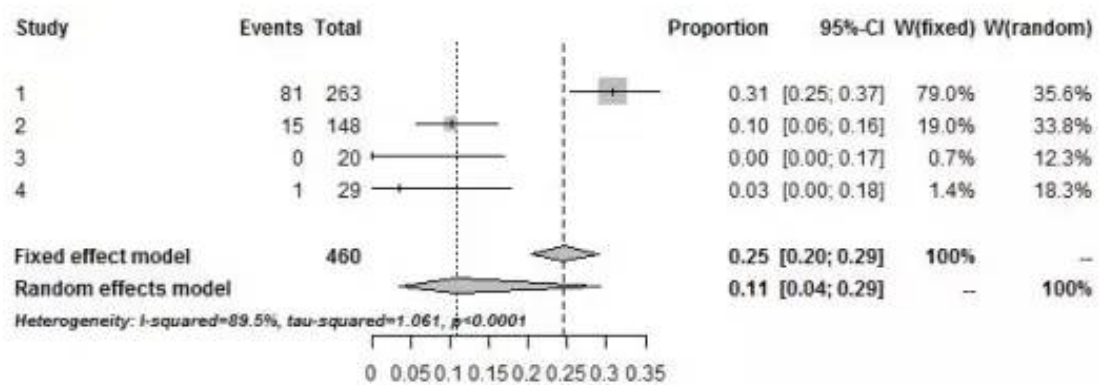
```
event <- c(81, 15, 0, 1)
```

```
n <- c(263, 148, 20, 29)
```

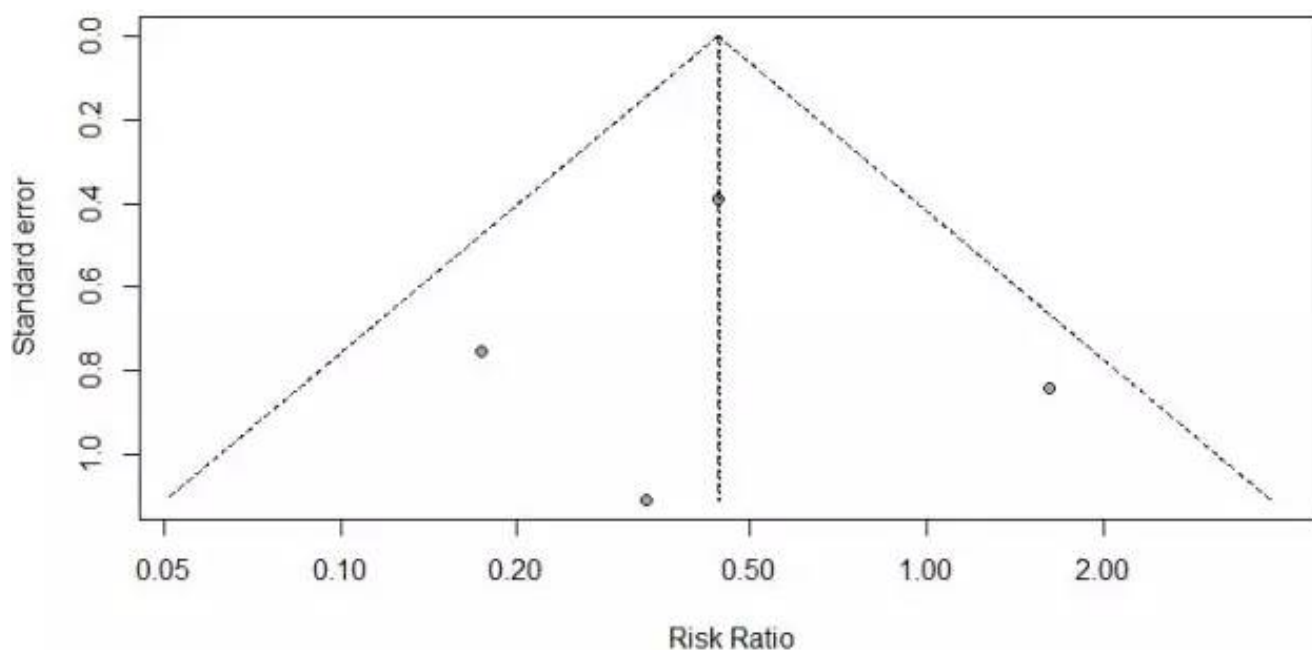
```
m1 <- metaprop(event, n, sm="PLOGIT")
```

```
forest(m1)
```

```
funnel(m1)
```



森林图效果：



8.连续变量效应指标的Meta分析R编程

对于连续性变量效应指标，同样可以用Meta包的metacont函数完成。

```
metacont(n.e, mean.e, sd.e, n.c, mean.c, sd.c,data= , sm=)
```

主要参数中，n.e, mean.e, sd.e分别表示试验组(暴露组)的样本量，均数，标准差;n.c, mean.c, sd.c表示对照组(非暴露组)的样本量，均数，标准差; sm标识汇总的效应指标，如smd,wmd等。

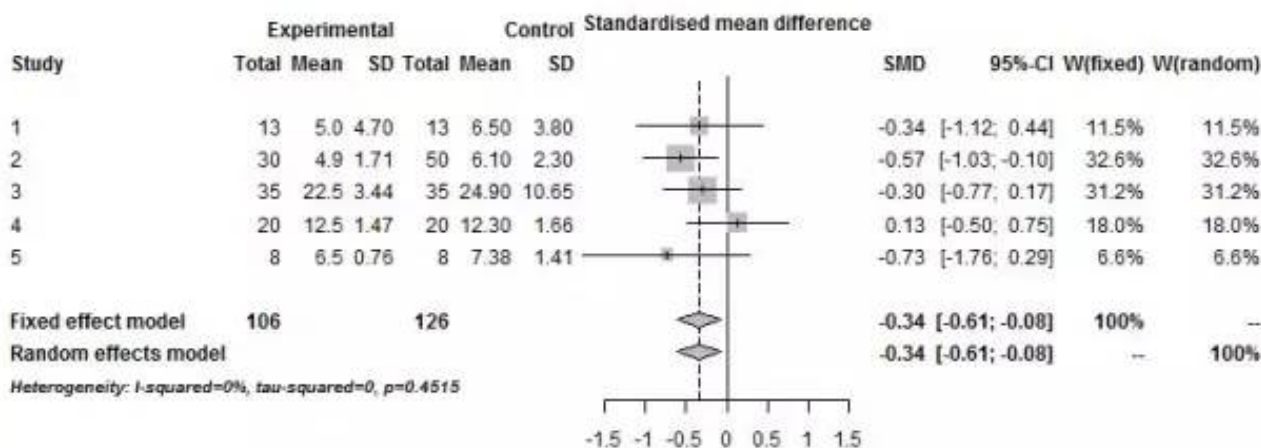
一个连续变量效应指标的Meta分析R编程示例：

```
data(Fleiss93cont)
```

```
meta2 <- metacont(n.e, mean.e, sd.e, n.c, mean.c, sd.c, data=Fleiss93cont, sm="SMD")
```

```
forest(meta2)
```

```
funnel(meta2)
```



9.分类变量效应指标的Meta分析R编程

对于分类变量效应指标，同样可以用Meta包的metabin函数完成。

```
metabin(event.e, n.e, event.c, n.c, data= , sm=)
```

主要参数中，event.e, n.e分别表示试验组(暴露组)的事件数和样本量;event.c, n.c,表示对照组(非暴露组)的事件数和样本量;sm表示汇总的效应指标，如RR,OR等。

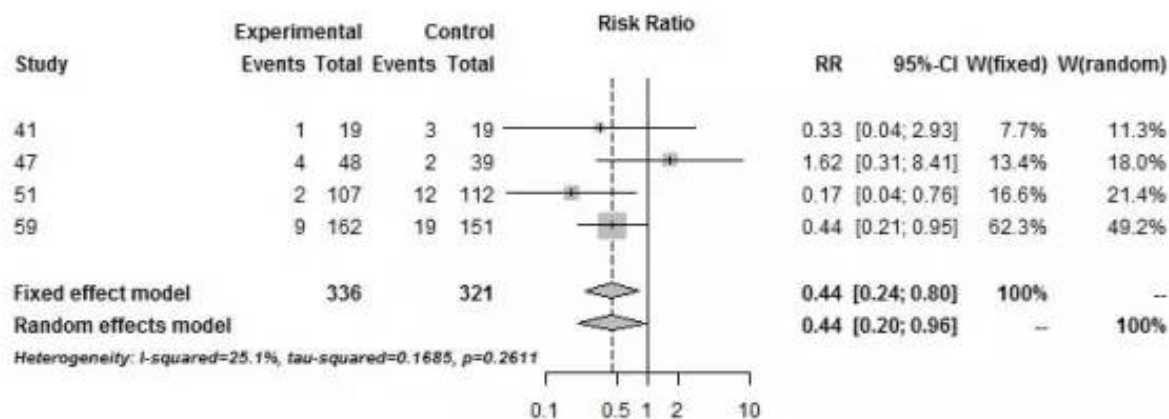
一个分类变量效应指标的Meta分析R编程示例：

```
data(Olkin95)

meta4<- metabin(event.e, n.e, event.c, n.c,data=Olkin95, subset=c(41,47,51,59),sm="RR",
method="I")

forest(meta4)

funnel(meta4)
```



10. Meta回归的R编程

meta回归的实现其实需要借助另一个包metafor, metafor的功能比meta包更丰富, 可以绘制各种丰富的图形(如前面提到的星状图、拉贝图、Begg's, Egger's检验图), 还可以拟合各种线性模型。

一个Meta回归的R编程的示例：

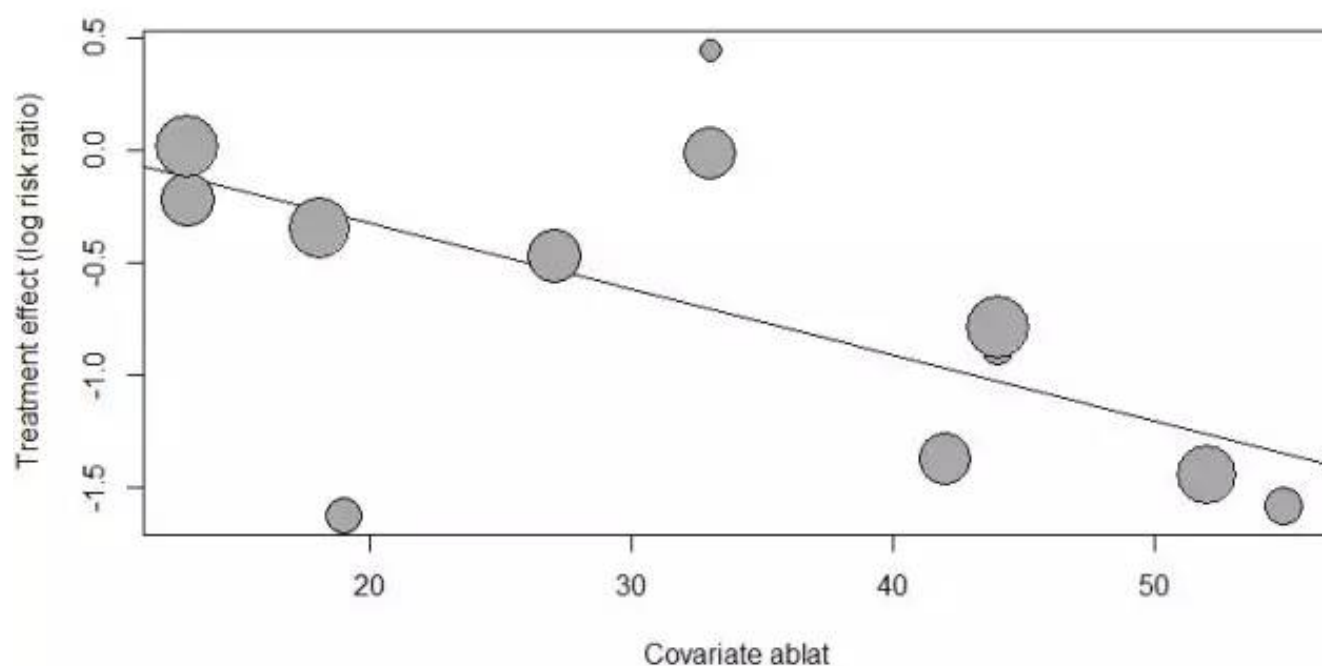
```
data(dat.colditz1994, package="metafor")

data10 <- dat.colditz1994

mh2 <- metabin(tpos, tpos+tneg, cpos, cpos+cneg, data=data10, studlab=paste(author, year))

mh2.mr <- metareg(mh2, ablat)

bubble(mh2.mr)
```



更多具体的函数和参数解释，可以查阅R软件包的帮助文档。此外，还有诊断试验的meta，网络meta本文未做说明。

最后，摘抄一句话，出自某位统计先哲，未找到出处，先搁这里，作为结尾。

如果你爱一个人，就让他做Meta分析，因为那里是天堂;如果你恨一个人，也让他做Meta分析，因为那里是地狱!

更多 论文写作 请访问 <https://www.iikx.com/news/article/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发