

·循证医学中的医学统计学问题·

在 R 软件中实现单个率的 Meta 分析

罗美玲, 谭红专, 周 权, 王莎亚, 蔡 畅, 郭亚伟, 沈 琳

(中南大学公共卫生学院流行病学与卫生统计学系, 长沙 410078)

[摘要] 目的 介绍在 R 软件中实现单个率的 Meta 分析方法。方法 在 R 软件中加载 Meta 分析程序包, 录入 Meta 分析数据, 用实例数据进行 Meta 分析。结果 单个率资料的 Meta 分析要求率的分布服从正态分布, R 软件提供了四种率的转换方法。对实例数据进行正态性转换后进行 Meta 分析, 其合并率及 95% 可信区间与原文一致。结论 R 软件可以实现对单个率研究结果的 Meta 分析, 功能强大, 实用性强。

[关键词] 率; Meta 分析; R 软件

[中图分类号] R195.1 **[文献标识码]** A **DOI:** 10.3969/j.issn.1671-5144.2013.03.014

Realizing the Meta-Analysis of Single Rate in R Software

LUO Mei-ling, TAN Hong-zhuan, ZHOU Quan, WANG Sha-ya, CAI Chang, GUO Ya-wei, SHEN Lin
(Department of Epidemiology and Health Statistics, School of Public Health, Central South University, Changsha 410078, China)

Abstract: **Objective** To introduce the method of meta-analysis for single rate in R software. **Method** We loaded meta-analysis package and inputted data, and then analyzed the example data in R software. **Results** Meta-analysis for single rate requested the distribution of rate was normal, and R software provided four transformation methods. The example results were consistent with the original literatures. **Conclusion** Meta-analysis for single rate could be realized in R software.

Key words: rate; meta-analysis; R software

Meta 分析主要用于汇总众多研究结果并进行定量分析^[1]。目前 Meta 分析的效应量有单个率(rate)、比值比(odds ratio, OR)、相对危险度(relative risk, RR)、均数(mean)等, 单个率的 Meta 分析是一种只提供了一组人群的总人数和事件发生人数, 基于原始研究为横断面的研究^[2]。各独立研究的效应量为率的 Meta 分析在国内外均有报道, 多个软件可以实现^[3-5]。国内 Meta 分析的软件主要是 Review Manager, 但该软件无法实现单个率的 Meta 分析^[6]。Stata 软件可进行单个率的 Meta 分

析, 但其计算繁琐, 操作较复杂且是收费的统计软件^[7]。Meta-analysis 软件是一款免费的软件且可以进行单个率的 Meta 分析, 但是其为菜单操作, 无法实现对原始率的转换^[8]。R 软件是一种共享的免费统计软件, 有专门的 Meta 分析程序包, 可以进行单个率的 Meta 分析, 而且提供了五种方法估算率, 研究者可以根据原始率的分布选择合适的方法^[9]。国内少有用 R 软件进行 Meta 分析的文献, 且没有介绍其具体操作步骤及适用条件^[10-11]。本文结合编程和 Meta 分析程序包, 以实例说明 R 软件在单个率 Meta 分析中的应用, 以期今后的 Meta 分析提供方法学指导。

1 R 软件介绍及 Meta 程序包导入

R 软件是一个自由、免费、源代码开放的软件, 可用于统计计算和统计制图的优秀工具^[12]。R 软

[作者简介] 罗美玲(1987-), 女, 土家族, 湖南湘西自治州人, 在读硕士研究生, 研究方向为分子流行病学。

[通讯作者] 谭红专, Tel:0731-88885835; E-mail:tanhz99@qq.com

件的最新版本可以从 CRAN (the Comprehensive R Archive Network) 下载, 网址为: <http://cran.r-project.org/>, 截止到 2012 年 10 月 30 日的最新版本是 R-2.15.2。

软件安装好后, 先在电脑中的任意盘 (如 D 盘) 建一个名为 Rworkplace 文件夹, 同时在桌面上的 R 软件图标上单击右键选择属性, 将“起始位置”修改为: D:\Rworkplace, 目的是使以后读文件方便。运行 R 软件, 点击“程序包”下拉菜单后, 再选“安装程序包”, 会出现下载数据源的镜像列表, 然后选择离自己距离最近的镜像, 如 China (Beijing1), 确定之后在出现的对话框中找到“meta”模块。最后加载 meta 程序包, 点击“程序包”图标后, 选“加载程序包”, 在列表中添加“meta”模块。

2 数据录入

以 Ford 等^[13]的研究在消化不良人群中肠易激综合征发生率的 Meta 分析数据为例 (见表 1)。建立 Excel 表, 按照图 1 中的格式录入原始数据, 录入原始研究的事件发生数 (event) 及样本量 (n), 然后另存为 CSV (逗号分隔) 文件, 放在之前建立的 D:\Rworkplace 文件夹中。

表 1 消化不良人群中肠易激综合征发生率的研究数据

No.	研究项目	事件发生数 (event)	消化不良 人数 (n)
1	Talley 1992	213	835
2	Holtmann 1994	44	180
3	Holtmann 1994	124	423
4	Talley 1994	200	919
5	Agreus 1995	372	1 154
6	Schlemper 1995	46	175
7	Schlemper 1995	80	473
8	Kennedy 1998	833	3 169
9	Talley 1998	92	730
10	Caballero-Plasencia 1999	63	264
11	Shah 2001	774	2 549
12	Curioso 2002	87	231
13	Hu 2002	304	1 649
14	Locke 2005	90	643
15	Lu 2005	561	2 018
16	Papatheoridis 2005	339	700
17	Perona 2005	47	70
18	Minocha 2006	247	990
19	Bolling-Sternevald 2008	397	1 001

	A	B	C	D
1	study	event	n	
2	Talley 1992	213	835	
3	Holtmann 1994	44	180	
4	Holtmann 1994	124	423	

图 1 数据录入格式

3 R 软件中的 Meta 分析函数

3.1 数据读入

可利用 read 函数读入之前建立好的 Excel 数据集, 相应的命令是: `rate<-read.csv("rate.csv")`

3.2 单个率的合并

可使用 metaprop 函数进行率的合并, R 软件中关于样本率的估计方法有五种, 根据样本率的分布决定使用哪种合并方法, 五种估计方法如下:

`sm="PRAW"`, Raw, i.e. untransformed, proportions (没有转换的原始率)。

`sm="PLN"`, Log transformation (对数转换)。

`sm="PLOGIT"`, Logit transformation (logit 转换)。

`sm="PAS"`, Arcsine transformation (反正弦转换)。

`sm="PFT"`, Freeman-Tukey Double arcsine transformation (Freeman-Tukey 双重反正弦转换)^[14]。

在进行 metaprop 分析之前, 对原始率及按四种估计方法进行转换后的率进行正态性检验, 根据检验结果选择接近正态分布的方法。transform 为对数据进行计算, p、log、logit、arcsin、dsrscin 表示分别按上述五种方法估计率的函数、write 为保存转换后的数据, shapiro.test 为正态性检验。命令如下:

```
transform(rate, p=event/n, log=log(event/n),
logit=log((event/n)/(1-event/n)), arcsin=asin(sqrt(
event/(n+1))), darcsin=0.5 * (asin(sqrt(event/
(n+1)))+asin(sqrt((event+1)/(n+1)))))->rate
write.csv(rate, file="rate.csv")
shapiro.test(rate$p); shapiro.test(rate$log);
shapiro.test(rate$logit); shapiro.test(rate$arcsin);
shapiro.test(rate$darcsin)
```

metaprop 函数用于计算各个独立研究的率及 95% 可信区间, 并按照率的分布选择合适的估计方法, 得到合并的率及 95% 可信区间。命令如下:

```
metarate<-metaprop(event, n, study, data=
rate, sm="PLOGIT", incr=0.5, allincr=TRUE,
addincr=FALSE, title="")
```

sm 表示估计样本率的方法, 根据正态性检验结果选择合适的估计方法。本例原始率的分布为

偏态,选取的是 PLOGIT 法估计原始率。

`incr=0.5` 表示当事件发生数为 0 时用 0.5 来替代。

`allincr=TRUE` 表示至少有一个研究的事件发生数为 0 时,所有事件发生数都加上 `incr` 数值。为 `FALSE` 时,只对事件发生数为 0 的研究加上 `incr` 数值。

`addincr=TRUE` 表示无论是否有事件发生数为

0 的研究,所有研究都加上 `incr` 数值。

`title=""` 表示这次 meta 分析的标题。

3.3 森林图的绘制

`forest` 是对计算的结果按森林图表示。命令为：`forest (metarate, digits=2)`。`digits` 表示保留的小数位数。合并的率及 95%可信区间与原文一致,结果见图 2。

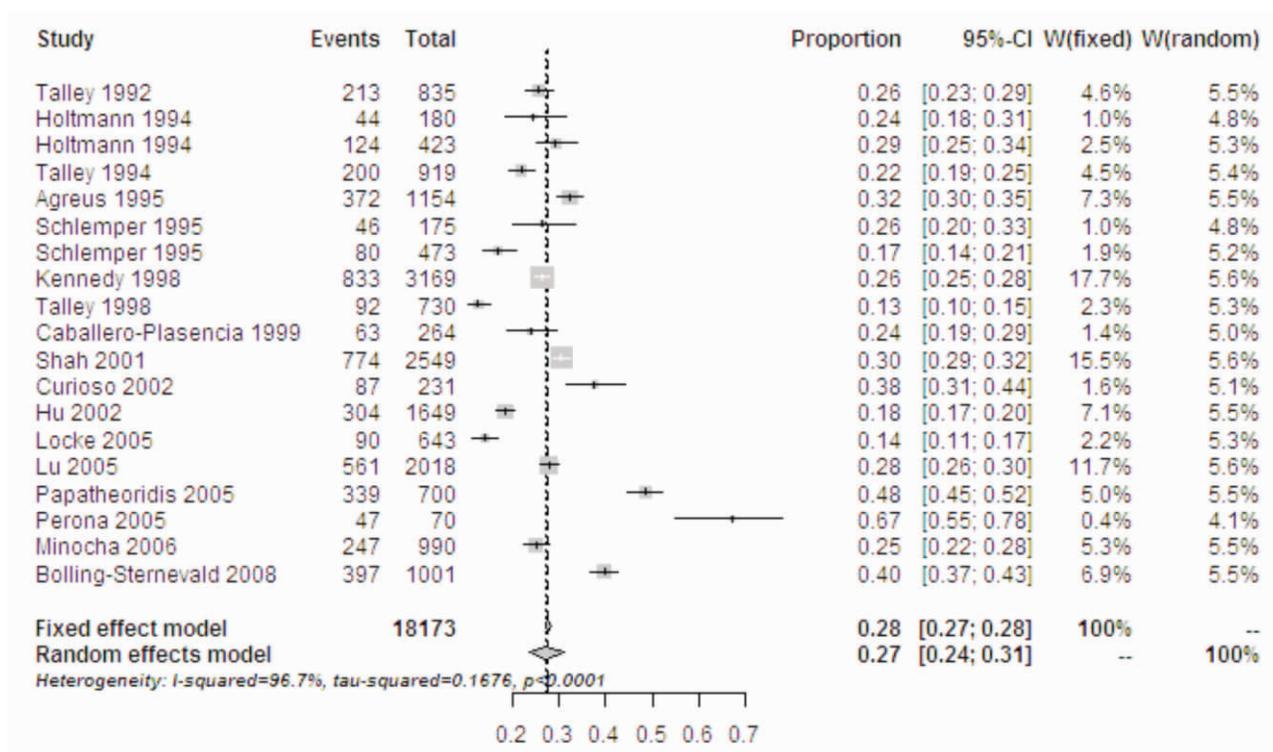


图 2 Meta 分析森林图

3.4 漏斗图的绘制

`funnel` 画漏斗图,是用于识别发表偏倚或其他偏倚的方法,根据图形的不对称程度判断 Meta 分析中偏倚的有无。

命令为:`funnel(metarate)`。结果见图 3。漏斗图是一种用主观定性的方法来判断有无偏倚,因此需要对漏斗图的不对称程度进行统计学检验。

3.5 漏斗图的不对称性检验

`metabias`是对漏斗图的不对称性进行统计学检验。

Egger 检验的命令为 `metabias (metarate, method="linreg")`,根据结果输出窗口 P 值大小,判断发表偏倚或其他偏倚是否具有统计学意义。输出 Egger 漏斗图的命令为 `metabias (metarate, method="linreg", plotit=TRUE, k.min=10)`。结果见图 4。Method 为 Egger 检验的方法;plotit 为画图

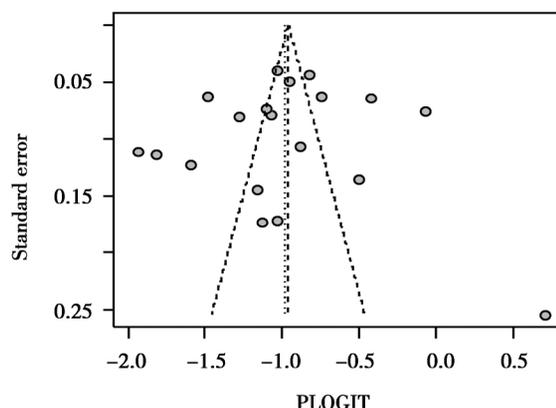


图 3 Meta 分析漏斗图

函数;k.min 为进行检验时所需最小的单个研究的数量,默认为 10,如果研究在 3~10 个之间可通过此函数进行调整。

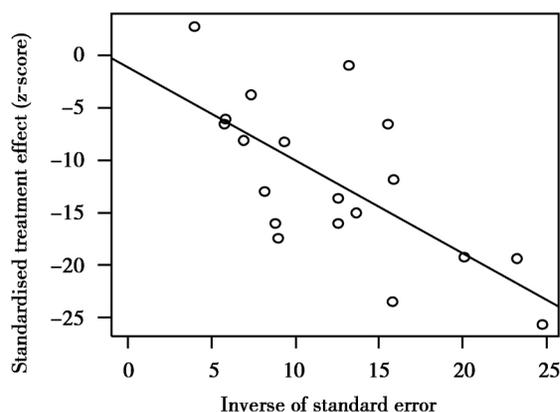


图4 Meta分析 Egger 漏斗图

3.6 敏感性分析

metainf 可以进行敏感性分析, 计算分别剔除每个入选研究后合并的 OR 值及 95% 可信区间。命令为: metainf(metarate)。

4 讨论

在科学研究中, 设立对照是一项基本原则, 如病例对照研究的病例组和对照组、队列研究中的暴露组和非暴露组, 临床随机对照试验的试验组和对照组。对这些研究进行 Meta 分析时合并的指标是两组的相对效应如 OR 值、RR 值或是绝对效应如危险度差值 (risk difference, RD), 然而在并未设立对照组如流行病学中的现况研究, 如要了解某种病毒在全国的一个总体感染率而又没有足够的时间或经费做全国性的调查时, 我们可以通过对现有文献报道的感染率进行 Meta 分析, 了解该病毒在全国的感染情况。

单个率可包括流行病学现况研究中的患病率、感染率, 临床试验的有效率, 药物试验中的不良反应发生率等, 只要收集到各个原始研究的样本量和事件发生数, 在 R 软件就可以用 metaprop 函数对单个率进行定量的 Meta 分析, 来计算合并的率及 95% 可信区间, 结合 funnel 函数, 可绘制漏斗图, 通过观察图形是否对称初步判断有无发表偏倚, 结合 metabin 可以绘制出 Begg 和 Egger 图, 对发表偏倚进行统计学检验。

考虑到不同类型的单个率的资料的分布可能会有不同的情况, R 软件给出了五种估计率的方法。如原始率不服从正态分布, 可经过转换使其服从或接近正态分布, 从而提高合并结果的可靠性。

本文选取的实例纳入文献数为 19, I^2 值为 96.7%, 各研究间存在异质性。由于单个率的 Meta 分析各原始文献为单个组别的率, 稳定性不同于具有两个组的研究, 因此在合并时统计学异质性一般会比较大会比较大, 这在国内外发表的关于单个率的 Meta 分析文章也可以看出^[11,15-16]。当异质性较大时, 首先要从专业性的角度对不同情况下的率进行亚组分析, 或通过敏感性分析来确定 Meta 分析的结果是否可靠。

本文结合实例, 介绍了在 R 软件中如何实现单个率的 Meta 分析。R 软件可以对 Meta 分析中多种计算方法进行快速简便且准确的统计分析^[17], 而且不断更新, 可以从网站上免费下载, 实用性强。R 软件在国内已逐渐得到应用, 其在 Meta 分析中的优势也将显现出来。

[参 考 文 献]

- [1] Glass GV. Primary, secondary and meta-analysis of research [J]. *Educ Res*, 1976, 6(5):3.
- [2] 曾宪涛, 冷卫东, 郭毅, 等. Meta 分析系列之一: Meta 分析的类型 [J]. *中国循证心血管医学杂志*, 2012, 4(1): 3-5.
- [3] 王佩鑫, 李宏田, 刘建蒙. 无对照二分类资料的 Meta 分析方法及 Stata 实现 [J]. *循证医学*, 2012, 12(1): 52-55.
- [4] Koskinen J, Löhönen J, Koponen H, et al. Prevalence of alcohol use disorders in schizophrenia—a systematic review and meta-analysis [J]. *Acta Psychiatr Scand*, 2009, 120(2): 85-96.
- [5] Batham A, Narula D, Toteja T, et al. Systematic review and meta-analysis of prevalence of hepatitis B in India [J]. *Indian Pediatr*, 2007, 44(9): 663-674.
- [6] 钟文昭, 吴一龙, 谷力加. Review manager (RevMan)——临床医生通向 Meta 分析的桥梁 [J]. *循证医学*, 2003, 3(4): 234-246.
- [7] 柏建岭, 钟文昭, 郑明华. Stata 在 Meta 分析中的应用 [J]. *循证医学*, 2007, 7(6): 363-368.
- [8] Wallace BC, Schmid CH, Lau J, et al. Meta-analyst: Software for meta-analysis of binary, continuous and diagnostic data [J]. *BMC Med Res Methodol*, 2009, 9(1): 80-91.
- [9] 张天嵩, 钟文昭. 实用循证医学方法学 2012 [M]. 长沙: 中南大学出版社, 2012: 363.
- [10] 杨国静, 孙乐平, 吴锋, 等. 50% 氯硝柳胺乙醇胺盐可湿性粉剂现场灭螺效果的 Meta 分析 [J]. *中国血吸虫病防治杂志*, 2010, 22(6): 579-582.
- [11] 何倩, 王静, 朱玉, 等. 中国内地居民 2000-2009 年被动吸烟率 Meta 分析 [J]. *中华流行病学杂志*, 2011, 32(2): 159-163.
- [12] Chongsuvivatwong V. Analysis of epidemiological data using R and Epicalc [M]. Songkhla, Thailand: Chanmuang Press, 2008.

(下转第 188 页)

[参 考 文 献]

- [1] Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2008 [J]. *CA Cancer J Clin*, 2008, 58(2):71-96.
- [2] Non-Small Cell Lung Cancer Collaborative Group. Chemotherapy in non-small cell lung cancer: A meta-analysis using updated data on individual patients from 52 randomized clinical trials [J]. *BMJ*, 1995, 311(7010):899-909.
- [3] Scagliotti GV, Fossati R, Torri V, et al. Randomized study of adjuvant chemotherapy for completely resected stage B or C non-small-cell lung cancer [J]. *J Natl Cancer Inst*, 2003, 95(19):1453-1461.
- [4] Waller D, Peake MD, Stephens RJ, et al. Chemotherapy for patients with non-small-cell lung cancer. The surgical setting of the Big Lung Trial [J]. *Eur J Cardio Thorac Surg*, 2004, 26(1):173-182.
- [5] Arriagada R, Bergman B, Dunant A, et al. Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer [J]. *N Engl J Med*, 2004, 350(4):351-360.
- [6] Le Chevalier T, Dunant A, Arriagada R, et al. Long-term results of the International Adjuvant Lung Cancer Trial (IALT) evaluating adjuvant Cisplatin-based chemotherapy in resected non-small-cell lung cancer (NSCLC) [J]. *J Clin Oncol*, 2010, 28(1):35-42.
- [7] Douillard JY, Rosell R, De Lena M, et al. Adjuvant Vinorelbine plus Cisplatin versus observation in patients with completely resected stage B- non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): A randomised controlled trial [J]. *Lancet Oncol*, 2006, 7(9):719-727.
- [8] Antoinette JW, Joan HS. Adjuvant treatment of non-small cell lung cancer: Improving cure rates with chemotherapy and other innovative therapies [J]. *ASCO Educational Book*, 2006, 410-414.
- [9] Pignon JP, Tribodet H, Scagliotti GV, et al. Lung adjuvant Cisplatin evaluation: A pooled analysis by the LACE Collaborative Group [J]. *J Clin Oncol*, 2008, 26(21):3552-3559.
- [10] Hermes A, Bergman B, Bremnes R, et al. Irinotecan plus Carboplatin versus oral Etoposide plus Carboplatin in extensive small-cell lung cancer: A randomized phase III trial [J]. *J Clin Oncol*, 2008, 26(26):4261-4267.
- [11] Strauss GM, Herndon JE, Maddaus MA, et al. Adjuvant Paclitaxel plus Carboplatin compared with observation in stage B non-small-cell lung cancer: CALGB 9633 with the Cancer and Leukemia Group B, Radiation Therapy Oncology Group, and North Central Cancer Treatment Group Study Groups [J]. *J Clin Oncol*, 2008, 26(31):5043-5051.
- [12] Butts CA, Ding K, Seymour L, et al. Randomized phase III trial of Vinorelbine plus Cisplatin compared with observation in completely resected stage B and C non-small-cell lung cancer: Updated survival analysis of JBR-10 [J]. *J Clin Oncol*, 2010, 28(1):29-34.
- [13] Mok TS, Wu YL, Thongprasert S, et al. Gefitinib or Carboplatin-Paclitaxel in pulmonary adenocarcinoma [J]. *N Engl J Med*, 2009, 361(10):947-957.
- [14] Maemondo M, Inoue A, Kobayashi K, et al. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR [J]. *N Engl J Med*, 2010, 362(25):2380-2388.
- [15] Mitsudomi T, Morita S, Yatabe Y, et al. Gefitinib versus Cisplatin plus Docetaxel in patients with non-small-cell lung cancer harbouring mutations of the epidermal growth factor receptor (WJTOG3405): An open label, randomised phase 3 trial [J]. *Lancet Oncol*, 2010, 11(2):121-128.
- [16] Goss GD, Lorimer I, Tsao MS, et al. A phase III randomized, double blind, placebo-controlled trial of the epidermal growth factor receptor inhibitor Gefitinib in completely resected stage B- non-small cell lung cancer (NSCLC): NCIC CTG BR.19 [J]. *J Clin Oncol*, 2010, 28(suppl 18):abstr LBA 7005.
- [17] Tsao MS, Sakurada A, Ding K, et al. Prognostic and predictive value of epidermal growth factor receptor tyrosine kinase domain mutation status and gene copy number for adjuvant chemotherapy in non-small cell lung cancer [J]. *J Thorac Oncol*, 2011, 6(1):139-147.

[收稿日期] 2012-05-04

(上接第 184 页)

- [13] Ford AC, Marwaha A, Lim A, et al. Systematic review and meta-analysis of the prevalence of irritable bowel syndrome in individuals with dyspepsia [J]. *Clin Gastroenterol Hepatol*, 2010, 8(5):401-409.
- [14] Freeman MF, Tukey JW. Transformations related to the angular and the square root [J]. *Ann Math Statist*, 1950, 21(4):607-611.
- [15] Ahmed B, Al-Khaffaf H. Prevalence of significant asymptomatic carotid artery disease in patients with peripheral vascular disease: A meta-analysis [J]. *Eur J Vasc Endovasc Surg*, 2009, 37(3):262-271.
- [16] Vallely A, Page A, Dias S, et al. The prevalence of sexually transmitted infections in Papua New Guinea: A systematic review and meta-analysis [J]. *PLoS ONE*, 2010, 5(12):e15586.
- [17] 蒋文瀚, 陈炳为, 郑建光, 等. R 语言 meta 包在 Meta 分析中的应用 [J]. *循证医学*, 2011, 11(5):305-309.

[收稿日期] 2012-11-07