

Meta分析中发表偏倚的Begg's检验、Egger's检验及Macaskill's检验的SAS程序实现

郑辉烈^{1,2} 王忠旭³ 王增珍^{1,*}

1. 华中科技大学同济医学院公共卫生学院流行病学与卫生统计系(武汉 430030)
2. 南昌大学医学院公共卫生学院流行病学与卫生统计系(南昌 330006); 3. 中国疾病预防控制中心(北京 100050)

摘要 Meta分析的结论易受各种偏倚的影响,“发表偏倚”是其中重要的偏倚之一。所以,评价所纳入的研究是否存在发表偏倚是Meta分析的一项重要的前期准备工作。Begg's检验、Egger's检验和Macaskill's检验是常见的用于定量评价Meta分析“发表偏倚”的方法,为了便于这些检验方法的应用,本文编写了基于SAS系统的这三个检验的程序,经实例数据分析,其运行结果与Stata软件完全相同。为研究人员提供了同时实现这三个检验的有效途径。

关键词 Meta分析;发表偏倚;Begg's检验;Egger's检验;Macaskill's检验

Applying the SAS Program to the Complete Begg's Test, the Egger's Test, and the Macaskill's Test for Publication Bias of Meta-analysis

ZHENG Hui-lie^{1,2}, WANG Zhong-xu³, WANG Zeng-zhen^{1,*}

1. Department of Epidemiology and Statistics, School of Public Health, Tongji Medical College of Huazhong University of Science and Technology, Wuhan 430030, China
2. Department of Epidemiology and Statistics, School of Public Health, Medical College of Nanchang University, Nanchang 330006, China
3. National Institute for Occupational Health and Poison Control, Chinese Center for Disease Control and Prevention, Beijing 100050, China

Abstract The conclusions of meta-analyses are susceptible to various of biases, and publication bias is one of such main bias. Therefore, Checking for evidence of publication bias should be undertaken routinely at the preliminary stage of a meta-analysis. Begg's test, Egger's test, and Macaskill's test are usually used to objectively identify publication bias in meta-analyses. In order to conveniently use these methods, the SAS program of these three tests was designed in this paper. In order test practical data, the fact that the output of this program of SAS software was consisted with the output of STATA software was validated. So, this program is an alternative way to do such hypothesis tests to identify the publication bias in meta-analyses.

Key words Meta-analysis; Publication bias; Begg's test; Egger's test; Macaskill's test

Meta分析是一种定量综合多个同类研究效应的文献综合研究方法,在文献综述时被广泛应用。但Meta分析的结论是否可信和有意义取决于是否能够客观而又全面地评价分析研究过程中的各种偏倚。一般情况下,Meta分析常常是以已发表的文献为综合研究的对象,而已发表的文献往往倾向于报告有统计学意义的结论,即无统计学意义的研究发表的可能性更小,这是Meta分析常见偏倚中的“发表偏倚(publication bias)”^[1-3]。所以,评价Meta分析中纳入的研究是否存在“发表偏倚”是进行Meta

分析的一项预备性的工作^[4]。评价Meta分析中的这种偏倚有多种方法,而应用较多的是可视化的“漏斗图(funnel plot)”法^[1,5]。所以,不仅Meta分析专用软件RevMan可以绘制该图,大部分统计软件如SAS、SPSS及Stata等也都能绘制“漏斗图”。可是“漏斗图”法作为一种通过定性判断图形是否对称来评价Meta分析是否存在“发表偏倚”的方法具有很大的主观性,不同的观察者可能得出不同的结论^[6]。鉴于此,Begg's检验^[7]、Egger's检验^[8]和Macaskill's检验^[9]等定量方法被应用于评价Meta分析的“发表偏倚”。Stata软件有专用的程序进行Begg's检验和Egger's检验,但Meta分析专用软件RevMan及常用统计软件SAS、SPSS等却没有进行上述检验的程序包。本文以实例数据编写了基于SAS系统的定量评价Meta分析“发表偏倚”的常见

基金资助: 国家科技部 973 课题(2002CB512910)

作者简介: 郑辉烈,男(1976年~),讲师,在读博士研究生,主要从事统计方法在医学中的应用研究。Email: zhenghuilie@yahoo.com.cn

* 通讯作者, Email: wzzh@mails.tjmu.edu.cn

方法 Egger's 检验、Begg's 检验和 Macaskill's 检验的程序,使这些检验能方便地在 SAS 系统软件中实现,供读者参考。

1 Begg's 检验^[7]、Egger's 检验^[8]和 Macaskill's 检验^[9]

Begg's 检验、Egger's 检验和 Macaskill's 检验是三种常见的应用于 Meta 分析定量评价其“发表偏倚”的统计检验方法。相对于通过定性判断图形是否对称来主观评价“发表偏倚”的“漏斗图”法,这三种检验方法所得结论更为客观可靠。

Begg's 检验计算较为复杂,通过判断标准化后的效应尺度 y^* 与方差 v 之间的相关性是否具有统计学意义来评价“发表偏倚”是否存在。此处的相关是指 y^* 与 v 的 Kendall's 相关^[10], y^* 的定义为: $y^* = (y_i - \bar{y}) / v_i^{1/2}$, 而此处 $\bar{y} = (\sum v_i^{-1} y_i) / \sum v_i^{-1}$, $v_i^* = v_i - (\sum v_i^{-1})^{-1}$, y 和 v 分别指未标准化的效应尺度及其方差,而 i, j 指 Meta 分析中的单个研究的序号。Kendall's 相关系数的取值介于 -1 到 +1 之间,越接近于 0,说明相关性越弱;等于 0,可认为无相关,即不存在“发表偏倚”,两变量间的相关系数是否等于 0 可通过假设检验作出判断。Begg's 检验计算 Kendall's 评分及其标准误,以其比值为检验统计量,作出 y^* 与 v 的 Kendall's 相关性是否存在的判断。Kendall's 评分的计算也较复杂,但该评分与 Kendall's τ 相关系数有数量关系,即 Kendall's 评分等于 $\tau \times k(k-1)/2$ (k 为 Meta 分析中包含的子研究的个数),可利用 SAS 系统中的相关分析程序模块先求出 Kendall's τ 相关系数后求得。Kendall's 评分标准误的计算公式为^[10]: $(k(k-1)(2k+5)/18)^{1/2}$ 。

Egger's 检验以标准化的效应尺度(如 $\ln(OR)$ 或 $SMD/\sqrt{\text{var}(SMD)}$) 为应变量 y 和以效应估计量的精度(如标准误的倒数)为自变量 x ,建立线性回归方程,如果回归方程的截距接近于 0,则认为发表偏倚较小,若截距等于 0,即认为不存在“发表偏倚”,截距是否等于零可通过假设检验来判断。 y 定义为:效应尺度 / 效应尺度的标准误; x 定义为: 1 / 效应尺度标准误。

Macaskill's 检验直接以研究效应尺度为变量 y 和研究的样本例数为自变量 n ,拟合加权线性回归方程,由于常见情况下漏斗图是以效应尺度为纵轴和以研究的样本例数为横轴的散点图,所以 Macaskill's 检验也被称为基于漏斗图的加权回归分析法。若斜率接近 0,并且无统计学意义,则可认为效应尺度与研究例数之间无数量依存关系,即 Meta 分析不存在

“发表偏倚”。权重为效应尺度的方差的倒数。

2 三种检验的 SAS 程序实现及实例分析

2.1 计数资料 Meta 分析的 Egger's 检验、Begg's 检验和 Macaskill's 检验

以 Shi 等^[11]一文为例说明以 OR 为效应尺度的 Meta 分析三种检验的 SAS 程序实现过程。该文纳入了 35 个病例对照研究,分析 CYP1A1 基因和 GSTM1 基因多态性与肺癌间的关系,其中有 20 个病例对照研究是关于 GSTM1 基因多态性的,本研究仅摘录了 GSTM1 基因多态性的数据作为实例,各病例对照研究的基本数据如表 1。

该 Meta 分析的“发表偏倚”的三种假设检验结果见图 1,以 0.1 为检验水准, Egger's 检验 ($P=0.058$) 提示本次纳入的研究可能存在“发表偏倚”,与原文的 Egger's 检验结论相同 ($P=0.058$); 同时, SAS 程序的 Begg's 检验 ($P=0.041$) 和 Macaskill's 检验 ($P=0.044$) 结果也显示纳入这些研究可能存在“发表偏倚”。三种检验结论一致,说明本次 Meta 分析研究的结果很可能受到了“发表偏倚”的影响,解释 Meta 分析的结论时要注意到这一点。经验证,本文 SAS 程序的 Begg's 检验和 Egger's 检验结果与 Stata

表 1 GSTM1 基因多态性与肺癌关系病例对照研究的基本数据^[11]

研究 第一作者	年份	肺癌		非肺癌	
		GSTM (-)	GSTM1 (+)	GSTM (-)	GSTM1 (+)
Gao Jianrui	1998	27	19	25	45
Hu Yiling	1998	34	25	29	30
Chen Senqing	1999	39	29	42	63
Lan Q	2000	82	40	60	62
Chen SQ	2001	56	50	39	67
Ying Chan	2002	43	13	65	34
Wang Jingwen	2003	97	67	90	91
Chang-Yeung M	2004	125	104	117	80
Luo Chenling	2004	45	18	24	23
Liang Geyu	2004	82	70	79	73
Li Weiyang	2004	127	90	95	105
Gu Yanfei	2004	101	79	102	122
Yang XR	2004	108	78	75	64
Li Dairong	2005	57	42	27	39
Zeng Mian	2005	56	35	51	40
Wang Na	2005	45	32	45	62
Ye Weiyun	2005	23	35	33	29
Li Ying	2005	59	39	61	77
Qian Biyun	2006	69	39	53	55
Wang Qiming	2006	40	16	19	23
合计	—	1 315	920	1 131	1 184

Begg's test					
Number of Studies	kendall's score	Std.Dev.of Score	Z	p> z	
20	64	30.822070015	2.0764341905	0.0378538118	
			Con.corrected 2.0439899062	0.0409545443	

Egger's test					
STD_EFF	coef.	Std.Err.	t	P> t	[95% Conf. Interval]
slope	-0.19673	0.30348	-0.65	0.5250	-0.834321615 0.4408665112
bias	2.26515	1.11680	2.03	0.0576	-0.081157668 4.6114674084

Macaskill's test					
EFF_SIZE	coef.	Std.Err.	t	P> t	[95% Conf. Interval]
Intercept	0.80067	0.19840	4.04	0.0008	0.3838524538 1.2174862336
sample_size	-0.00144	0.00066386	-2.16	0.0443	-0.002830181 -0.000040733

图 1 GSTM1基因多态性与肺癌关系Meta分析发表偏倚检验结果

软件运行结果一致。

三种检验的 SAS 程序:

**** 第一部分: 数据输入及效应尺度及其标准误计算
程序 ****/

/* case_n1: 病例组阳性数; case_n0: 病例组阴性数;
con_n1: 对照组阳性数; con_n0: 对照组阴性数; study_size:
研究个数。*/

option linesize=120;

%let study_size=20;

data eff_sizeset;

input case_n1 case_n0 con_n1 con_n0@@;

k=_n_;

/*ln(OR) 及其标准误的计算, 若选择其它效应尺度, 可
在此作相应修改 */

sample_size=case_n0+case_n1+con_n0+con_n1;

or=(case_n1*con_n0/(case_n0*con_n1));

eff_size=log(or);

var_eff_size=1/case_n1+1/case_n0+1/con_n0+1/con_n1;

se_eff_size=sqrt(var_eff_size);

/* 产生 egger's 检验用应变量 y 和自变量 x*/

egger_y=eff_size/se_eff_size;

egger_x=1/se_eff_size;

/* 产生效应尺度 (ln(OR)) 的方差倒数、效应尺度与其
方差倒数之积 */

inv_v=1/var_eff_size;

eff_size_inv_v=eff_size*inv_v;

cards;

27 19 25 45 34 25 29 30

39 29 42 63 82 40 60 62

56 50 39 67 43 13 65 34

97 67 90 91 125 104 117 80

45 18 24 23 82 70 79 73

127 90 95 105 101 79 102 122

108 78 75 64 57 42 27 39

56 35 51 40 45 32 45 62

23 35 33 29 59 39 61 77

69 39 53 55 40 16 19 23

;

run;

***** 第一部分结束 *****/

**** 第二部分: 三种检验的实现程序及结果报表输出
程序 ****/

/* egger's 检验程序 */

ods listing close;

ods output parameterestimates=egg1

(keep=Variable Estimate StdErr tValue Probt

rename=(variable=STD_EFF Estimate=coefficient));

proc reg;

model egger_y=egger_x;

run;

ods listing;

data eggset;

set egg1;

if _n_=1 then STD_EFF='bias';else STD_EFF='slope';

CIL_95=coefficient-tinv(0.975,&study_size-2)*stderr;

CIU_95=coefficient+tinv(0.975,&study_size-2)*stderr;

run;

data egg2(where=(STD_EFF="slope")) egg3(where=(STD
_EFF="bias"));

set eggset;

```

run;
/* 产生 egger's 检验输出数据集 eggset*/
data eggset;
merge egg2 egg3(rename=(STD_EFF=STD_EFF1
coefficient=coefficient1 StdErr=StdErr1 Probt=Probt1
tValue=tValue1 CIL_95=CIL_951 CIU_95=CIU_951))
; run;
quit;
/*begg's 检验程序 */
/* 产生 ln(OR) 的方差倒数之和 (sum_inv_v)*/
data beg1;
set eff_sizeset;
sum0+inv_v;
if _n_=&study_size then sum=sum0;else sum=0;run;
proc sort; by descending k;
run;
data beg2;
set beg1;
sum_inv_v+sum;
run;
/* 产生 ln(OR) 和其方差倒数之积的总和 (eff_size_sum_
inv_v)*/
data beg3;
set beg2;
sum1+eff_size_inv_v;
if _n_=&study_size then sum=sum1;else sum=0;run;
proc sort; by k;
run;
data beg4;drop sum sum0 sum1;
set beg3;
sum_eff_size_inv_v+sum;
run;
/* 产生 ln(OR) 的平均值、ln(OR)- 的方差 v* 及 y* */
data eff_sizeset1;
set beg4;
y_mean=sum_eff_size_inv_v/sum_inv_v;
v_asterisk=var_eff_size-1/sum_inv_v;
y_asterisk=(eff_size-y_mean)/sqrt(v_asterisk);
run;
/* 产生 kendall's 系数  $\tau$  */
proc corr data=eff_sizeset1 kendall out=kenda(keep=var_
_eff_size) noprint;
var y_asterisk var_eff_size; run;
proc transpose data=kenda out=kendall3(keep=col3 col4
rename=(col3=n col4=tau));

```

```

/* 产生 kendall's 得分及其标准误、Begg's 检验 z 值及双
侧 P 值 */
data kendall;set kendall3;
kend_score=n*(n-1)*tau/2;
se_kend_score=sqrt(n*(n-1)*(2*n+5)/18);
z=kend_score/se_kend_score;
con_correct_z=((abs(kend_score)-1)*(abs(kend_score)/
kend_score)/se_kend_score);
p_z=2*(1-probnorm(abs(z)));p_con_correct_z=2*(1-
probnorm(abs(con_correct_z)));
run;
/* 产生 macaskill's 检验程序 */
ods listing close;
ods output parameterestimates=mac1(keep=Variable
Estimate StdErr tValue Probt
rename=(variable=STD_EFF Estimate=coefficient));
proc reg data=eff_sizeset1;
model eff_size=sample_size;
weight inv_v;
run;
ods listing;
data macset;
set mac1;
CIL_95=coefficient-tinv(0.975,&study_size-2)*StdErr;
CIU_95=coefficient+tinv(0.975,&study_size-2)*StdErr; run;
data mac2(where=(STD_EFF="Intercept")) mac3(where=
(STD_EFF="sample_size"));
set macset;
run;
/* 产生 Macaskill's 检验输出数据集 Macaskill*/
data macset;
merge mac2(rename=(STD_EFF=STD_EFF2
coefficient=coefficient2 StdErr=StdErr2 Probt=Probt2
tValue=tValue2 CIL_95=CIL_952 CIU_95=CIU_952))
mac3(rename=(STD_EFF=STD_EFF3 coefficient=coefficient3
StdErr=StdErr3 Probt=Probt3 tValue=tValue3
CIL_95=CIL_953 CIU_95=CIU_953))
; run;
/* 产生三种检验的输出数据集 (begg_egger_mac) 及三
种检验的结果报表输出程序 */
data begg_egger_mac;
merge kendall eggset macset;
file 'c:\Publicationbias.txt' print n=ps notitle;
put
#1 @34 "Begg's test"

```

```
#2 @2 99*'-1'
#3 @2 "Number of Studies" @22 "kendall's score" @39 "Std.
Dev.of Score" @72 "Z" @89 "p>|z|"
#4 @2 99*'-1'
#5 @2 "&study_size" @22 kend_score @39 se_kend_score
@72 z @89 p_z
#6 @57 "Con.corrected" @72 con_correct_z @89
p_con_correct_z
#7 @2 99*'-1'
#9 @34 "Egger's test"
#10 @2 78*'-1'
#11 @2 'STD_EFF' @12 "coef." @22 'Std.Err.' @33 't' @42
'P>|t|' @56 '[95% Conf. Interval]'
#12 @2 78*'-1'
#13 @2 STD_EFF @11 coefficient @22 StdErr @33 tValue
@42 Probt @52 CIL_95 @67 CIU_95
#14 @2 STD_EFF1 @11 coefficient1 @22 StdErr1 @33
tValue1 @42 Probt1 @52 CIL_951 @67 CIU_951
```

```
#15 @2 78*'-1'
#17 @34 "Macaskill's test"
#18 @2 80*'-1'
#19 @2 'EFF_SIZE' @15 "coef." @25 "Std.Err." @36 't' @45
'P>|t|' @59 "[95% Conf. Interval]"
#20 @2 80*'-1'
#21 @2 STD_EFF2 @15 coefficient2 @25 StdErr2 @36
tValue2 @45 Probt2 @55 CIL_952 @70 CIU_952
#22 @2 STD_EFF3 @15 coefficient3 @25 StdErr3 @36
tValue3 @45 Probt3 @55 CIL_953 @70 CIU_953
#23 @2 80*'-1';
run;
dm "inc 'c:\Publicationbias.txt";run; quit;
/***** 第二部分结束 *****/
```

2.2 计量资料Meta分析的Egger's检验、Begg's检验和Macaskill's检验

以 Zhuo^[12] 一文为例,说明以均数之差为效应

表 2 大豆蛋白摄入量与血清低密度脂蛋白胆固醇 ($\bar{X}\pm SD$) 关系的随机对照临床试验的基本数据^[12]

研究 第一作者	年份	高大蛋白摄入组		低大豆蛋白摄入组	
		n	LDC-C (mmol/L)	n	LDC-C (mmol/L)
Crouse JR ¹	1999	15	4.22±0.47	12	4.53±0.44
Crouse JR ²	1999	15	3.83±0.31	33	3.80±0.80
Mackey R	2000	25	4.71±0.73	41	4.18±0.70
Merz-Demlow BE	2000	13	2.14±0.18	22	4.82±0.91
Gardner CD	2001	31	3.50±0.50	24	4.78±0.96
Wangen KE	2001	18	3.01±0.21	15	3.85±0.52
Jenkins DJ	2002	41	4.14±0.70	20	3.80±0.59
Lichtenstein AH ¹	2002	22	4.70±1.12	13	2.31±0.18
Lichtenstein AH ²	2002	20	3.83±0.43	24	2.87±0.49
Steinberg FM	2003	24	2.86±0.49	18	3.22±0.21

Begg's test					
Number of Studies	kendall's score	Std.Dev.of Score	Z	p> z	
10	-17	11.180939887	-1.520526225	0.1283787729	
			Con.corrected	-1.431083506 0.152406284	

Egger's test						
STD_EFF	coef.	Std.Err.	t	P> t	[95% Conf. Interval]	
slope	0.67454	0.52336	1.29	0.2335	-0.532324405	1.8814068373
bias	-3.13515	1.72638	-1.82	0.1069	-7.116177714	0.8458803996

Macaskill's test						
EFF_SIZE	coef.	Std.Err.	t	P> t	[95% Conf. Interval]	
Intercept	-0.60705	0.32574	-1.86	0.0994	-1.358218396	0.1441163236
sample_size	0.00681	0.00599	1.14	0.2886	-0.00700871	0.0206324828

图 2 大豆蛋白摄入量与LDL-C关系的Meta分析发表偏倚检验结果

尺度的 Meta 分析三种检验的 SAS 程序实现过程。该文纳入 8 个随机对照试验,进行 Meta 分析以阐明大豆蛋白摄入量降低血清低密度脂蛋白胆固醇的效果^[12]。由于其中两个试验中同时包括了高胆固醇血症试验对象和正常试验对象,而其余 6 个试验只有其中一种类型的试验对象,所以,该文中实际包括 10 个比较研究。从原文中摘录了各研究的第一作者、年份、例数、均数和标准差等基本数据如表 2。

该研究原文未对“发表偏倚”进行评价,是一明显瑕疵。运行本文 SAS 程序,三种检验的结果见图 2,以 0.1 为检验水准, Begg's 检验 ($P=0.152$)、 Egger's 检验 ($P=0.107$) 和 Macaskill's 检验 ($P=0.289$) 结果均提示,该研究没有“发表偏倚”存在。经验证,本文 SAS 程序的 Begg's 检验和 Egger's 检验结果与 Stata 软件运行结果一致。三种检验的结论一致说明本研究不存在“发表偏倚”,为利用 Meta 分析结论提供了有力的支持。

三种检验的 SAS 程序:

```
**** 第一部分:数据输入及效应尺度及其标准误计算程序 ****/
```

```
/* n1: 样本 1 例数 ;mean1: 样本 1 均数 ;sd1 样本 1 标准差 ;n2: 样本 2 例数 ;mean2: 样本 2 均数 ;sd2 样本 2 标准差 ;study_size: 研究个数 */
```

```
option linesize=120;
%let study_size=10;
data eff_sizeset;
input n1 mean1 sd1 n2 mean2 sd2@@;
k=_n_;
sample_size=n1+n2;
/* 标准均数差及其 cohen 标准误的计算,若选择其它效应尺度,可在此作相应修改 */
md=mean1-mean2;
eff_size=md/sqrt(((sd1**2)*(n1-1)+(sd2**2)*(n2-1))/(sample_size-2));
var_eff_size=sample_size/n1/n2+(eff_size**2)/2/(sample_size-2);
se_eff_size=sqrt(var_eff_size);
/* 产生 egger's 检验用应变量 y 和自变量 x */
egger_y=eff_size/se_eff_size;
egger_x=1/se_eff_size;
/* 产生效应尺度(SMD)的方差倒数、效应尺度与其方差倒数之积 */
inv_v=1/var_eff_size;
```

```
eff_size_inv_v=eff_size*inv_v;
cards;
15 4.22 0.47 12 4.53 0.44
15 3.83 0.31 15 3.85 0.52
25 4.71 0.73 24 4.78 0.96
13 2.14 0.18 13 2.31 0.18
31 3.50 0.50 33 3.80 0.80
18 3.01 0.21 18 3.22 0.21
41 4.14 0.70 41 4.18 0.70
22 4.70 1.12 22 4.82 0.91
20 3.83 0.43 20 3.80 0.59
24 2.86 0.49 24 2.87 0.49
run;
/***** 第一部分结束 *****/
```

```
**** 第二部分:三种检验的实现程序及结果报表输出程序 ****/
```

与计数资料三种检验实现程序的第二部分完全相同,此处略。

```
**** 第二部分结束 *****/
```

3 结语

Meta 分析作为一种应用于文献综述研究的定量研究方法,在总结健康相关问题证据并为公共卫生、医疗卫生服务等提供决策依据方面发挥着越来越重要的作用。但 Meta 分析的结论易受到各种偏倚的影响。Meta 分析不可能纳入待研究问题的所有研究(即已发表和未发表的研究),往往只纳入了已发表的研究结果,而已发表和未发表的研究的特征可能会有明显的区别,这会导致 Meta 分析的结论出现“发表偏倚”^[1,3,13]。所以,利用 Meta 分析结论作决策之前,评价 Meta 分析的发表偏倚是一项必不可少的工作^[4]。

“漏斗图”法简便、直观,已被广泛应用于评价 Meta 分析研究中的“发表偏倚”是否存在。漏斗图常常是以效应大小作为横坐标,样本含量为纵坐标所作的散点图,若纳入的研究无发表偏倚,则图形呈现倒置的漏斗形;若漏斗图不对称或不完整,则可能存在发表偏倚^[5,14]。所以,“漏斗图”法通过直观判断漏斗图形的对称性来鉴定“发表偏倚”是否存在,是一种定性的评价方法。对同一漏斗图形,不同的观察者可能会得出不同的鉴定结论。基于此,定量通过显著性检验来评价“发表偏倚”是否存在的方法,如 Begg's 检验、Egger's 检验和 Macaskill's 检验等被应用于 Meta 分析的前期准备工作。Begg's

检验以标准化的效应估计量和效应估计量的方差是否存在 Kendall's 相关来判定“发表偏倚”是否存在,若相关性存在,认为“发表偏倚”存在;反之,不存在^[7]。Egger's 检验以标准化的效应估计量为应变量和效应估计量的精度(标准误的倒数)为自变量建立简单线性回归模型,经假设检验,若截距不等于零,即认为“发表偏倚”存在;反之,不能认为偏倚存在^[8]。Macaskill's 检验以效应估计量为应变量和样本例数为自变量,并以效应估计量的方差倒数为权重建立加权线性回归模型,经假设检验,若回归系数不等于零,即认为“发表偏倚”存在;反之,不存在^[9]。Macaskill 等^[15]通过模拟数据研究发现,这三种检验在实际的 Meta 分析所包含的研究个数条件下,检验效能均较低,但相对来说,基于漏斗图的加权线性回归分析法——Macaskill's 检验法为最优选择。而 Hayashino 等^[16]通过实际数据比较发现,相对于 Macaskill's 检验, Egger's 检验和 Begg's 检验效能更大;这与 Macaskill 等的研究结论不一致,作者的解释是由于模拟数据与实际数据的差别造成的。我们认为, Macaskill 等的结论更为可信,因为该结论是通过设定不同的参数经过大量的模拟数据得到每个方法平均的检验效能值。但不管如何,它们的研究结论给我们选择应用这三种检验提供了参考。我们认为,在评价 Meta 分析的“发表偏倚”时,同时选择使用这三种检验,也许是一种较好的选择,结合应用“漏斗图”法,若三种结论一致,表明判断结论更为可信;若不一致,则提示要慎重判断是否存在“发表偏倚”。现常见的统计软件到目前为止都没有发现同时完成这三个检验的模块,需要研究者自己编写程序完成。本研究以实例数据为基础,只需读者输入计数或计量资料的基本数据,实现了在 SAS 系统中进行这三种检验,而 Stata 软件只提供了 Egger's 检验和 Begg's 检验。同时,程序还实现了统计结果报表输出功能,提高了程序输出结果的可阅读性。经实例数据分析发现, Egger's 检验和 Begg's 检验的结果与 Stata 软件缺省设置条件下的结果一致,而且读者可以在程序的说明之处作简单的修改来实现 Stata 软件其它设置条件下的检验。在 SAS

软件中成功实现这三种检验,方便了这些检验在评价“发表偏倚”中的应用,为读者实现这些检验提供了另一选择。

参 考 文 献

- 1 Begg CB, Berlin JA. Publication bias: A problem in interpreting medical data. *Journal of the Royal Statistical Society, Series A*, 1988, 151(3): 419-463.
- 2 Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ*, 1994, 309(6964): 1286-1291.
- 3 Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Education and Prevention*, 1997, 9(1 Suppl): 15-21.
- 4 Begg CB. Publication bias. In *The Handbook of Research Synthesis*, Cooper H, Hedges LV (eds). Russell Sage Foundation: New York, 1994, Chapter 25.
- 5 Light RJ, Pillemer DB. *Summing Up: The Science of Reviewing Research*. Harvard University Press: Cambridge MA, 1984.
- 6 Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol*, 1994, 140: 290-296.
- 7 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*, 1994, 50: 1088-1101.
- 8 Egger M, Smith GD, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 1997, 315: 629-634.
- 9 Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med*, 2001, 20: 641-654.
- 10 Kendall MG. A New Measure of Rank Correlation. *Biometrika*, 1938, 30(1-2): 81-93.
- 11 Shi XQ, Zhou SH, Wang ZX, et al. CYP1A1 and GSTM1 polymorphisms and lung cancer risk in Chinese populations: A meta-analysis. *Lung Cancer*, 2008, 59: 155-163.
- 12 Zhuo XG, Melby MK, Watanabe S. Soy Isoflavone Intake Lowers Serum LDL Cholesterol: A Meta-Analysis of 8 Randomized Controlled Trials in Humans. *J Nutr*, 2004, 134: 2395-2400.
- 13 Easterbrook PJ, Berlin JA, Gopalan R, et al. Publication bias in clinical research. *Lancet*, 1991, 337: 867-872.
- 14 Sutton AJ, Duval SJ, Tweedie RL, et al. Empirical assessment of effect of publication bias on metaanalyses. *BMJ*, 2000, 320: 1574-1577.
- 15 Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. *Stat Med*, 2001, 20: 641-654.
- 16 Hayashino Y, Noguchi Y, Fukui T. Systematic evaluation and comparison of statistical tests for publication bias. *Journal of Epidemiology*, 2005, 15(6): 235-243.

收稿日期: 2008-10-29

本文编辑: 杜亮