

# 18 Meta-analysis in Stata<sup>TM</sup>

JONATHAN A C STERNE, MICHAEL J BRADBURN,  
MATTHIAS EGGER

## Summary points

- Stata<sup>TM</sup> is a general-purpose, command-line driven, programmable statistical package.
- A comprehensive set of user-written commands is freely available for meta-analysis.
- Meta-analysis of studies with binary (relative risk, odds ratio, risk difference) or continuous outcomes (difference in means, standardised difference in means) can be performed.
- All the commonly used fixed effect (inverse variance method, Mantel-Haenszel method and Peto's method) and random effect (DerSimonian and Laird) models are available.
- An influence analysis, in which the meta-analysis estimates are computed omitting one study at a time, can be performed.
- Forest plots, funnel plots and L'Abbé plots can be drawn and statistical tests for funnel plot asymmetry can be computed.
- Meta-regression models can be used to analyse associations between treatment effect and study characteristics.

We reviewed a number of computer software packages that may be used to perform a meta-analysis in Chapter 17. In this chapter we show in detail how to use the statistical package Stata both to perform a meta-analysis and to examine the data in more detail. This will include looking at the accumulation of evidence in cumulative meta-analysis, using graphical and statistical techniques to look for evidence of bias, and using meta-regression to investigate possible sources of heterogeneity.

## Getting started

Stata is a general-purpose, command-line driven, programmable statistical package in which commands to perform several meta-analytic methods

All data sets described in this Chapter are available from the book's website:  
<[www.systematicreviews.com](http://www.systematicreviews.com)>.

### **Box 18.1 Downloading and installing user-written meta-analysis commands**

As a first step we recommend that you make sure that your installation is up-to-date by typing **update all** in the command window. Stata will automatically connect to www.stata.com and update the core package. It will also download brief descriptions of all user-written commands published in the Stata Technical Bulletin. Those relating to meta-analysis can be displayed by typing **search meta**. The most convenient way to install user-written commands is from within Stata. Go into the “Help” menu and click on the “STB and User-Written Programs” option. Now click on <http://www.stata.com> and then on stb (for Stata Technical Bulletins). The meta-analysis routines described in this chapter can then be downloaded as follows:

<i>Click on...</i>	<i>... then click on</i>	<i>to install commands</i>
stb45	sbe24.1	<b>metan, funnel, labbe</b>
stb43	sbe16.2	<b>meta</b>
stb42	sbe22	<b>metacum</b>
stb56	sbe26.1	<b>metainf</b>
stb58	sbe19.3	<b>metabias</b>
stb42	sbe23	<b>metareg</b>

Note that these are the latest versions as of December 2000 and you should check whether updated versions or new commands have become available (**update all, search meta**).

are available. Throughout this chapter, Stata commands appear in **bold font**, and are followed by the Stata output that they produce. Users should note that the commands documented here do not form part of the “core” Stata package, but are all user-written “add-ons” which are freely available on the internet. In order to perform meta-analyses in Stata, these routines need to be installed on your computer by downloading the relevant files from the Stata web site ([www.stata.com](http://www.stata.com)). See Box 18.1 for detailed instructions on how to do this.

We do not attempt to provide a full description of the commands: interested readers are referred to help files for the commands, and to the relevant articles in the *Stata Technical Bulletin* (STB, see reference list). To display the help file, type **help** followed by the command (for example **help metan**) or go into the “Help” menu and click on the “Stata command...” option. Bound books containing reprints of a year’s Stata

Technical Bulletin articles are also available and are free to university libraries. The articles referred to in this chapter are available in STB reprints volumes 7: (STB 38 to STB 42) and 8 (STB 43 to 48). The Stata website gives details of how to obtain these. All the output shown in this chapter was obtained using Stata version 6. Finally, we assume that the data have already been entered into Stata.

## Commands to perform a standard meta-analysis

### Example 1: intravenous streptokinase in myocardial infarction

The following table gives data from 22 randomised controlled trials of streptokinase in the prevention of death following myocardial infarction.<sup>1-3</sup>

Table 18.1

Trial number	Trial name	Publication year	Intervention group		Control group	
			Deaths	Total	Deaths	Total
1	Fletcher	1959	1	12	4	11
2	Dewar	1963	4	21	7	21
3	1st European	1969	20	83	15	84
4	Heikinheimo	1971	22	219	17	207
5	Italian	1971	19	164	18	157
6	2nd European	1971	69	373	94	357
7	2nd Frankfurt	1973	13	102	29	104
8	1st Australian	1973	26	264	32	253
9	NHLBI SMIT	1974	7	53	3	54
10	Valere	1975	11	49	9	42
11	Frank	1975	6	55	6	53
12	UK Collaborative	1976	48	302	52	293
13	Klein	1976	4	14	1	9
14	Austrian	1977	37	352	65	376
15	Lasierra	1977	1	13	3	11
16	N German	1977	63	249	51	234
17	Witchitz	1977	5	32	5	26
18	2nd Australian	1977	25	112	31	118
19	3rd European	1977	25	156	50	159
20	ISAM	1986	54	859	63	882
21	GISSI-1	1986	628	5860	758	5852
22	ISIS-2	1988	791	8592	1029	8595

These data were saved in Stata dataset **strepto.dta** which is available from the book's website (<http://www.systematicreviews.com>). We can list the variables contained in the dataset, with their descriptions (variable labels) by using the **describe** command:

## SYSTEMATIC REVIEWS IN HEALTH CARE

### **describe**

```
Contains data from strepto.dta
obs: 22                               Streptokinase after MI
vars: 7
size: 638 (99.7% of memory free)
```

```
1. trial      byte   %8.0g    Trial number
2. trialnam   str14  %14s     Trial name
3. year       int    %8.0g    Year of publication
4. pop1       int    %12.0g   Treated population
5. deaths1    int    %12.0g   Treated deaths
6. pop0       int    %12.0g   Control population
7. deaths0    int    %12.0g   Control deaths
```

Sorted by: trial

### **The metan command**

The **metan** command<sup>4</sup> provides methods for the meta-analysis of studies with two groups. With binary data the effect measure can be the difference between proportions (sometimes called the risk difference or absolute risk reduction), the ratio of two proportions (risk ratio or relative risk), or the odds ratio. With continuous data both observed differences in means or standardised differences in means can be used. For both binary and continuous data either fixed effects or random effects models can be fitted.

For analysis of trials with binary outcomes, the command requires variables containing the number of individuals who did and did not experience disease events, in intervention and control groups. Using the streptokinase data, the variables required can be created as follows:

```
generate alive1=pop1-deaths1
generate alive0=pop0-deaths0
```

In the following, we use the **metan** command to perform a meta-analysis on relative risks, derive the summary estimate using Mantel-Haenszel methods, and produce a forest plot. The options (following the comma) that we use are:

```
rr                      perform calculations using relative risks
xlab(.1,1,10)           label the x-axis
label(namevar=trialnam) label the output and vertical axis of the
                         graph with the trial name. The trial year
                         may also be added by specifying
                         yearvar=year.
```

META-ANALYSIS IN STATA

Display the help file for a complete list of options. The command and output in our analysis are as follows (note that all commands are typed on one line although they may be printed on two):

```
metan deaths1 alive1 deaths0 alive0, rr xlab(.1,1,10)
label(namevar=trialnam)
```

Study	RR	[95% Conf Interval]	%	Weight
Fletcher	.229167	.030012	1.74987	.177945
Dewar	.571429	.196152	1.66468	.298428
1st European	1.3494	.742948	2.45088	.63566
Heikinheimo	1.22321	.668816	2.23714	.74517
Italian	1.0105	.551044	1.85305	.784121
2nd European	.702555	.533782	.924693	4.0953
2nd Frankfurt	.457066	.252241	.828213	1.22434
1st Australian	.778646	.478015	1.26835	1.39327
NHLBI SMIT	2.37736	.648992	8.70863	.126702
Valere	1.04762	.480916	2.28212	.413208
Frank	.963636	.33158	2.80052	.260532
UK Collab	.895568	.626146	1.28092	2.25043
Klein	2.57143	.339414	19.4813	.051901
Austrian	.608042	.417252	.886071	2.67976
Lasierra	.282051	.033993	2.3403	.138556
N German	1.16088	.840283	1.60379	2.24179
Witchitz	.8125	.26341	2.5062	.235214
2nd Australian	.849654	.536885	1.34463	1.28713
3rd European	.509615	.33275	.78049	2.11133
ISAM	.880093	.619496	1.25031	2.65037
GISSI-1	.827365	.749108	.913797	32.3376
ISIS-2	.768976	.704392	.839481	43.8613
M-H pooled RR	.79876	.754618	.845484	

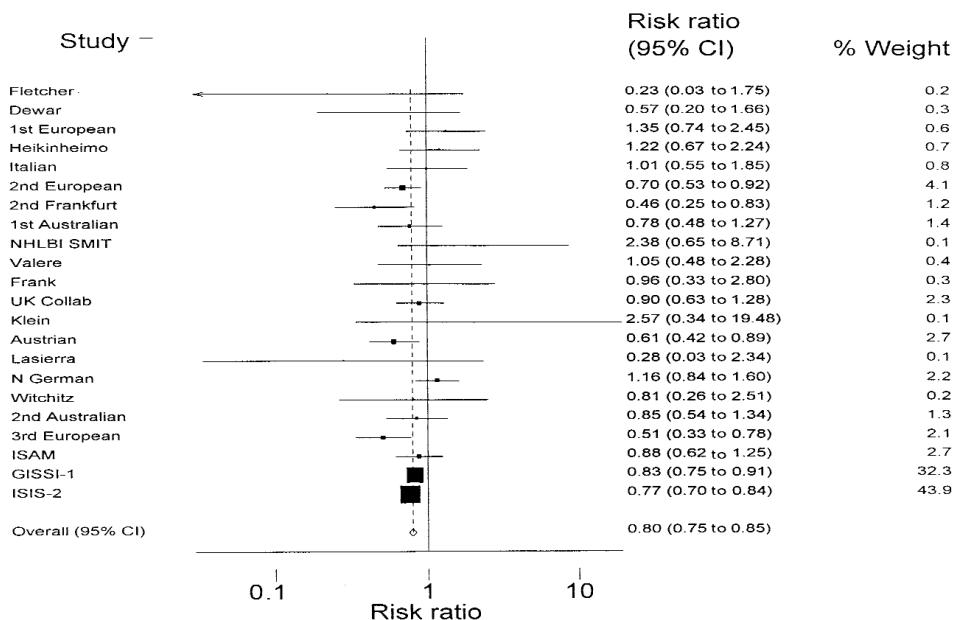
Heterogeneity chi-squared = 30.41 (d.f. = 21) p = 0.084  
Test of RR=1 : z= 7.75 p = 0.000

The output shows, for each study, the treatment effect (here, the relative risk) together with the corresponding 95% confidence interval and the percentage weight contributed to the overall meta-analysis. The summary (pooled) treatment effect (with 95% CI and P value) and the heterogeneity test are also shown. By default, new variables containing the treatment

## SYSTEMATIC REVIEWS IN HEALTH CARE

effect size, its standard error, the 95% CI and study weights and sample sizes are added to the dataset.

The **metan** command also automatically produces a forest plot (see Chapter 2). In a forest plot the contribution of each study to the meta-analysis (its weight) is represented by the area of a box whose centre represents the size of the treatment effect estimated from that study (point estimate). The confidence interval for the treatment effect from each study is also shown. The summary treatment effect is shown by the middle of a diamond whose left and right extremes represent the corresponding confidence interval.



Both the output and the graph show that there is a clear effect of streptokinase in protecting against death following myocardial infarction. The meta-analysis is dominated by the large GISSI-1<sup>2</sup> and ISIS-2<sup>3</sup> trials which contribute 76·2% of the weight in this analysis. If required, the text showing the weights or treatment effects may be omitted from the graph (options **nowt** and **nostats**, respectively). The **metan** command will perform all the commonly used fixed effects (inverse variance method, Mantel-Haenszel method and Peto's method) and random effects (DerSimonian and Laird) analyses. These methods are described in Chapter 15. Commands **tabbe** to draw L'Abbé plots (see Chapters 8 and 10) and **funnel** to draw funnel plots (see Chapter 11) are also included.

**The meta command**

The **meta** command<sup>5–7</sup> uses inverse-variance weighting to calculate fixed and random effects summary estimates, and, optionally, to produce a forest plot. The main difference in using the **meta** command (compared to the **metan** command) is that we require variables containing the effect estimate and its corresponding standard error for each study. Commands **metacum**, **metainf**, **metabias** and **metareg** (described later in this chapter) also require these input variables. Here we re-analyse the streptokinase data to demonstrate **meta**, this time considering the outcome on the odds ratio scale. For odds ratios or risk ratios, the **meta** command works on the log scale. So, to produce a summary odds ratio we need to calculate the log of the ratio and its corresponding standard error for each study. This is straightforward for the odds ratio. The log odds ratio is calculated as

```
generate logor=log((deaths1/alive1)/(deaths0/alive0))
```

and its standard error, using Woolf's method, as

```
generate selogor=sqrt((1/deaths1)+(1/alive1)+  
(1/deaths0)+(1/alive0))
```

Chapter 15 gives this formula, together with the standard errors of the risk ratio and other commonly used treatment effect estimates. The output can be converted back to the odds ratio scale using the **eform** option to exponentiate the odds ratios and their confidence intervals. Other options used in our analysis are:

<b>graph(f)</b>	display a forest plot using a fixed-effects summary estimate. Specifying <b>graph(r)</b> changes this to a random-effects estimate
<b>cline</b>	draw a broken vertical line at the combined estimate
<b>xlab(.1,1,10)</b>	label the <i>x</i> -axis at odds ratios 0·1, 1 and 10
<b>xline(1)</b>	draw a vertical line at 1
<b>id(trialnam)</b>	label the vertical axis with the trial name contained in variable <b>trialnam</b>
<b>b2title(Odds ratio)</b>	label the <i>x</i> -axis with the text “Odds ratio”.
<b>print</b>	output the effect estimates, 95% CI and weights for each study

SYSTEMATIC REVIEWS IN HEALTH CARE

The command and output are as follows:

```
meta logor selogor, eform graph(f) cline xline(1)
xlab(.1,1,10) id(trialnam) b2title(Odds ratio) print
```

Meta-analysis (exponential form)

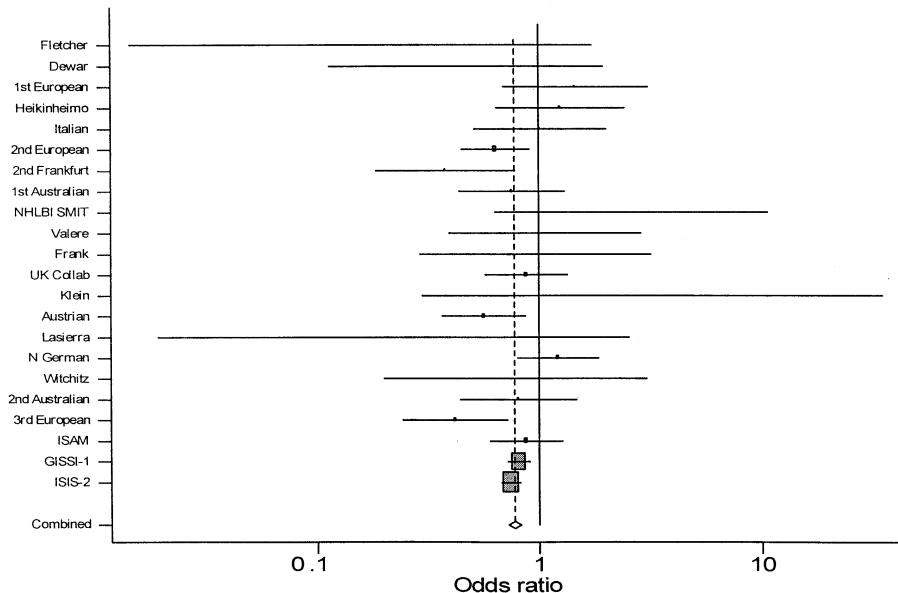
Method	Pooled		95% CI		Asymptotic		No. of studies
	Est	Lower	Upper	z_value	p_value		
Fixed	0.774	0.725	0.826	-7.711	0.000		22
Random	0.782	0.693	0.884	-3.942	0.000		

Test for heterogeneity: Q= 31.498 on 21 degrees of freedom (p= 0.066)

Moment-based estimate of between studies variance = 0.017

Study	Weights		Study Est	95% CI	
	Fixed	Random		Lower	Upper
Fletcher	0.67	0.67	0.16	0.01	1.73
Dewar	1.91	1.85	0.47	0.11	1.94
1st European	6.80	6.10	1.46	0.69	3.10
Heikinheimo	8.72	7.61	1.25	0.64	2.42
Italian	8.18	7.19	1.01	0.51	2.01
2nd European	31.03	20.39	0.64	0.45	0.90
2nd Frankfurt	7.35	6.54	0.38	0.18	0.78
1st Australian	12.75	10.50	0.75	0.44	1.31
NHLBI SMIT	1.93	1.87	2.59	0.63	10.60
Valere	3.87	3.63	1.06	0.39	2.88
Frank	2.67	2.55	0.96	0.29	3.19
UK Collab	20.77	15.39	0.88	0.57	1.35
Klein	0.68	0.67	3.20	0.30	34.59
Austrian	20.49	15.24	0.56	0.36	0.87
Lasierra	0.65	0.64	0.22	0.02	2.53
N German	21.59	15.84	1.22	0.80	1.85
Witchitz	2.06	1.99	0.78	0.20	3.04
2nd Australian	10.50	8.92	0.81	0.44	1.48
3rd European	13.02	10.68	0.42	0.24	0.72
ISAM	27.13	18.63	0.87	0.60	1.27
GISSI-1	303.12	49.69	0.81	0.72	0.90
ISIS-2	400.58	51.76	0.75	0.68	0.82

## META-ANALYSIS IN STATA



Note that **meta** performs both fixed and random effects analyses by default and the tabular output includes the weights from both analyses. It is clear that the smaller studies are given relatively more weight in the random effects analysis than with the fixed effect model. Because the **meta** command requires only the estimated treatment effect and its standard error, it will be particularly useful in meta-analyses of studies in which the treatment effect is not derived from the standard  $2 \times 2$  table. Examples might include crossover trials, or survival trials, when the treatment effect might be measured by the hazard ratio derived from Cox regression.

### Example 2: intravenous magnesium in acute myocardial infarction

The following table gives data from 16 randomised controlled trials of intravenous magnesium in the prevention of death following myocardial infarction. These trials are a well-known example where the results of a meta-analysis<sup>8</sup> were contradicted by a single large trial (ISIS-4)<sup>9-11</sup> (see also Chapters 3 and 11).

SYSTEMATIC REVIEWS IN HEALTH CARE

Table 18.2

Trial number	Trial name	Publication year	Intervention group		Control group	
			Deaths	Total	Deaths	Total
1	Morton	1984	1	40	2	36
2	Rasmussen	1986	9	135	23	135
3	Smith	1986	2	200	7	200
4	Abraham	1987	1	48	1	46
5	Feldstedt	1988	10	150	8	148
6	Schechter	1989	1	59	9	56
7	Ceremuzynski	1989	1	25	3	23
8	Bertschat	1989	0	22	1	21
9	Singh	1990	6	76	11	75
10	Pereira	1990	1	27	7	27
11	Schechter 1	1991	2	89	12	80
12	Golf	1991	5	23	13	33
13	Thogersen	1991	4	130	8	122
14	LIMIT-2	1992	90	1159	118	1157
15	Schechter 2	1995	4	107	17	108
16	ISIS-4	1995	2216	29 011	2103	29 039

These data were saved in Stata dataset **magnes.dta**.

**describe**

Contains data from magnes.dta

obs: 16 Magnesium and CHD  
vars: 7

---

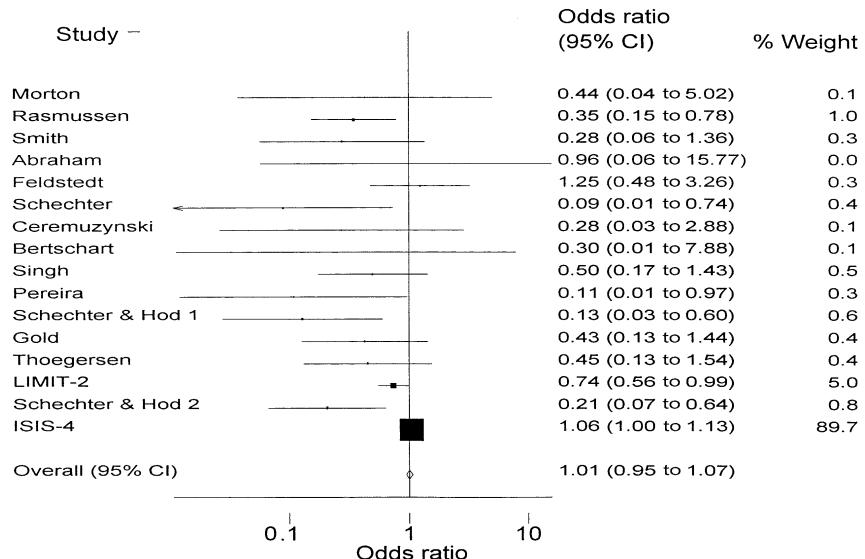
1.	trial	int	%8.0g	Trial number
2.	trialnam	str12	%12s	Trial name
3.	year	int	%8.0g	Year of publication
4.	tot1	long	%12.0g	Total in magnesium group
5.	dead1	double	%12.0g	Deaths in magnesium group
6.	tot0	long	%12.0g	Total in control group
7.	dead0	long	%12.0g	Deaths in control group

---

Sorted by: trial

The discrepancy between the results of the ISIS-4 trial and the earlier trials can be seen clearly in the graph produced by the **metan** command. Note that because the ISIS-4 trial provides 89·7% of the total weight in the meta-analysis, the overall (summary) estimate using fixed-effects analysis is very similar to the estimate from the ISIS-4 trial alone.

## META-ANALYSIS IN STATA



### Dealing with zero cells

When one arm of a study contains no events – or, equally, all events – we have what is termed a “zero cell” in the  $2 \times 2$  table. Zero cells create problems in the computation of ratio measures of treatment effect, and the standard error of either difference or ratio measures. For trial number 8 (Bertschart), there were no deaths in the intervention group, so that the estimated odds ratio is zero and the standard error cannot be estimated. A common way to deal with this problem is to add 0.5 to each cell of the  $2 \times 2$  table for the trial. If there are no events in either the intervention or control arms of the trial, however, then any measure of effect summarised as a ratio is undefined, and unless the absolute (risk difference) scale is used instead, the trial has to be discarded from the meta-analysis.

The **metan** command deals with the problem automatically, by adding 0.5 to all cells of the  $2 \times 2$  table before analysis. For the commands which require summary statistics to be calculated (**meta**, **metacum**, **metainf**, **metabias** and **metareg**) it is necessary to do this, and to drop trials with no events or in which all subjects experienced events, before calculating the treatment effect and standard error.

To drop trials with no events or all events:

```
drop if dead1==0&dead0==0
drop if dead1==tot1&dead0==tot0
```

SYSTEMATIC REVIEWS IN HEALTH CARE

To add 0·5 to the  $2 \times 2$  table where necessary:

```
gen trzero=0
replace trzero=1 if
dead1==0|dead0==0|dead1==tot1|dead0==tot0
(1 real change made)
replace dead1=dead1+0.5 if trzero==1
(1 real change made)
replace dead0=dead0+0.5 if trzero==1
(1 real change made)
replace tot1=tot1+1 if trzero==1
(1 real change made)
replace tot0=tot0+1 if trzero==1
(1 real change made)
```

To derive summary statistics needed for meta-analysis:

```
generate alive0=tot0-dead0
generate alive1=tot1-dead1
generate logor=log((dead1/alive1)/(dead0/alive0))
generate
selogor=sqrt((1/dead1)+(1/alive1)+(1/dead0)+(1/alive0))
```

To use the **meta** command to perform a meta-analysis:

```
meta logor selogor, eform id(trialnam) print
```

Meta-analysis (exponential form)

Method	Pooled Est	95% CI Lower	Asymptotic z_value	p_value	No. of studies
Fixed	1.015	0.956	1.077	0.484	0.629
Random	0.483	0.329	0.710	-3.706	0.000

Test for heterogeneity: Q= 47.059 on 15 degrees of freedom (p= 0.000)

Moment-based estimate of between studies variance = 0.224

## META-ANALYSIS IN STATA

Study	Weights		Study Est	95% CI	
	Fixed	Random		Lower	Upper
Morto	0.64	0.56	0.44	0.04	5.02
Rasmussen	5.83	2.53	0.35	0.15	0.78
Smith	1.53	1.14	0.28	0.06	1.36
Abraham	0.49	0.44	0.96	0.06	15.77
Feldstedt	4.18	2.16	1.25	0.48	3.26
Schechter	0.87	0.73	0.09	0.01	0.74
Ceremuzynski	0.70	0.61	0.28	0.03	2.88
Bertschart	0.36	0.34	0.30	0.01	7.88
Singh	3.48	1.96	0.50	0.17	1.43
Pereira	0.81	0.69	0.11	0.01	0.97
Schechter & Hod 1	1.64	1.20	0.13	0.03	0.60
Gold	2.61	1.65	0.43	0.13	1.44
Thoegersen	2.55	1.62	0.45	0.13	1.54
LIMIT-2	46.55	4.08	0.74	0.56	0.99
Schechter & Hod 2	3.03	1.81	0.21	0.07	0.64
ISIS-4	998.78	4.45	1.06	1.00	1.13

Note the dramatic difference between the fixed and random effects summary estimates, which arises because the studies are weighted much more equally in the random effects analysis. Also, the test of heterogeneity is highly significant. We will return to this example later.

## Cumulative meta-analysis

The **metacum** command<sup>12</sup> performs and graphs cumulative meta-analyses,<sup>13,14</sup> in which the cumulative evidence at the time each study was published is calculated. This command also requires variables containing the effect estimate and its corresponding standard error for each study (see above). To perform a cumulative meta-analysis of the streptokinase trials, we first create a character variable of length 20 containing both trial name and year, and then sort by year:

```
gen str21 trnamyr=trialnam+" "+string(year)+""
sort year
```

The options for the **metacum** command are similar to those for the **meta** command, except:

SYSTEMATIC REVIEWS IN HEALTH CARE

**effect(f)** perform all calculations using fixed-effects meta-analysis. Specifying **effect(r)** changes this to a random-effects estimate  
**graph** produce a cumulative meta-analysis graph

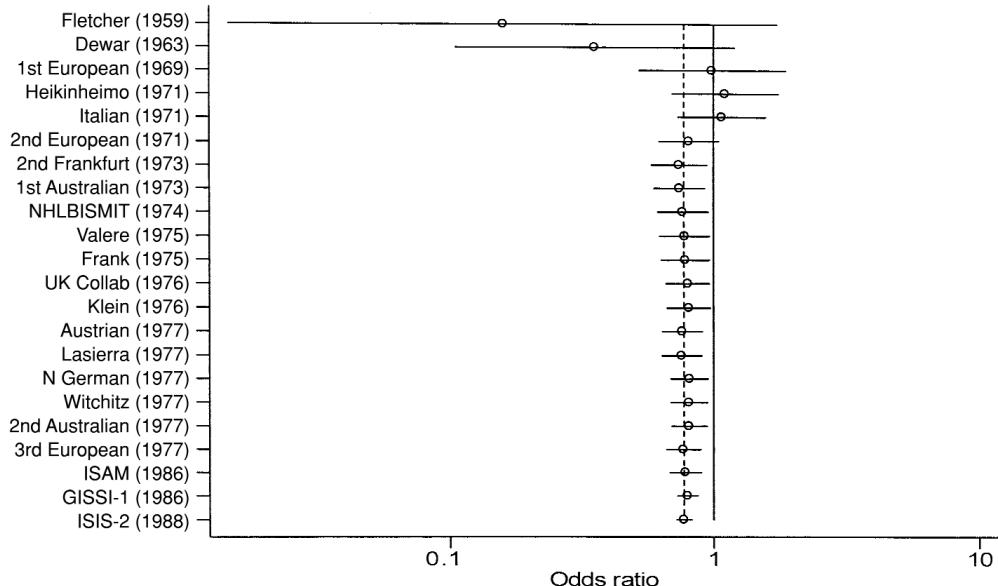
The command and output are as follows:

```
metacum logor selogor, effect(f) eform graph cline
xline(1) xlabel(.1,.1,10) id(trnamyr) b2title(Odds ratio)
```

Cumulative fixed-effects meta-analysis of 22 studies  
(exponential form)

Trial	Cumulative		95% CI		z	P value
	estimate	Lower	Upper			
Fletcher (1959)	0.159	0.015	1.732	-1.509	0.131	
Dewar (1963)	0.355	0.105	1.200	-1.667	0.096	
1st European (1969)	0.989	0.522	1.875	-0.034	0.973	
Heikinheimo (1971)	1.106	0.698	1.753	0.430	0.667	
Italian (1971)	1.076	0.734	1.577	0.376	0.707	
2nd European (1971)	0.809	0.624	1.048	-1.607	0.108	
2nd Frankfurt (1973)	0.742	0.581	0.946	-2.403	0.016	
1st Australian (1973)	0.744	0.595	0.929	-2.604	0.009	
NHLBI SMIT (1974)	0.767	0.615	0.955	-2.366	0.018	
Valere (1975)	0.778	0.628	0.965	-2.285	0.022	
Frank (1975)	0.783	0.634	0.968	-2.262	0.024	
UK Collab (1976)	0.801	0.662	0.968	-2.296	0.022	
Klein (1976)	0.808	0.668	0.976	-2.213	0.027	
Austrian (1977)	0.762	0.641	0.906	-3.072	0.002	
Lasierra (1977)	0.757	0.637	0.900	-3.150	0.002	
N German (1977)	0.811	0.691	0.951	-2.571	0.010	
Witchitz (1977)	0.810	0.691	0.950	-2.596	0.009	
2nd Australian (1977)	0.810	0.695	0.945	-2.688	0.007	
3rd European (1977)	0.771	0.665	0.894	-3.448	0.001	
ISAM (1986)	0.784	0.683	0.899	-3.470	0.001	
GISSI-1 (1986)	0.797	0.731	0.870	-5.092	0.000	
ISIS-2 (1988)	0.774	0.725	0.826	-7.711	0.000	

## META-ANALYSIS IN STATA



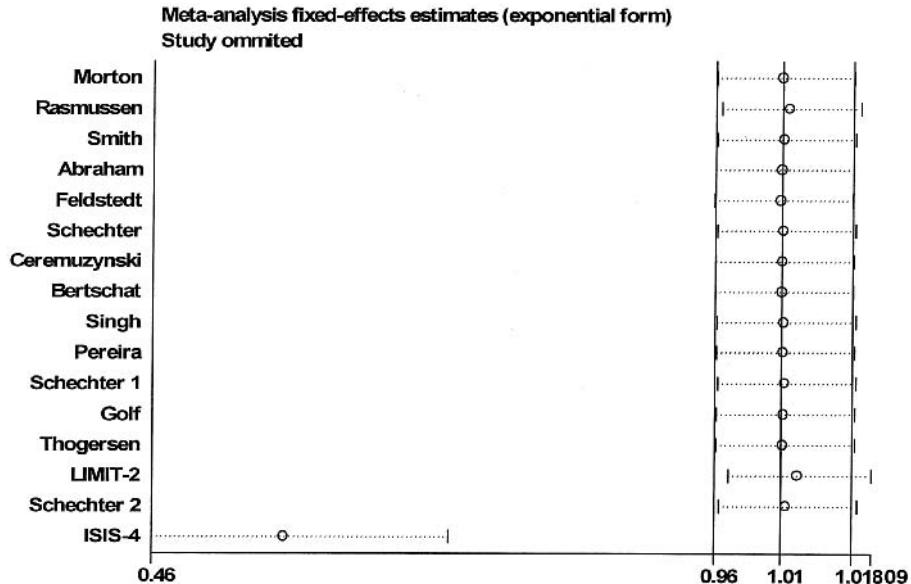
By the late 1970s, there was clear evidence that streptokinase prevented death following myocardial infarction. However it was not used routinely until the late 1980s, when the results of the large GISSI-1 and ISIS-2 trials became known (see Chapter 1). The cumulative meta-analysis plot makes it clear that although these trials reduced the confidence interval for the summary estimate, they did not change the estimated degree of protection.

### Examining the influence of individual studies

The influence of individual studies on the summary effect estimate may be displayed using the **metainf** command.<sup>15</sup> This command performs an influence analysis, in which the meta-analysis estimates are computed omitting one study at a time. The syntax for **metainf** is the same as that for the **meta** command. By default, fixed-effects analyses are displayed. Let's perform this analysis for the magnesium data:

```
metainf logor selogor, eform id (trialnam)
```

## SYSTEMATIC REVIEWS IN HEALTH CARE



The label above the vertical axis indicates that the treatment effect estimate (here, log odds ratio) has been exponentiated. The meta-analysis is dominated by the ISIS-4 study, so omission of other studies makes little or no difference. If ISIS-4 is omitted then there appears to be a clear effect of magnesium in preventing death after myocardial infarction.

### Funnel plots and tests for funnel plot asymmetry

The **metabias** command<sup>16,17</sup> performs the tests for funnel-plot asymmetry proposed by Begg and Mazumdar<sup>18</sup> and by Egger *et al.*<sup>11</sup> (see Chapter 11). If the **graph** option is specified the command will produce either a plot of standardized effect against precision<sup>11</sup> (**graph(egger)**) or a funnel plot (**graph(begg)**). For the magnesium data there is clear evidence of funnel plot asymmetry if the ISIS-4 trial is included. It is of more interest to know if there was evidence of bias *before* the results of the ISIS-4 trial were known. Therefore in the following analysis we omit the ISIS-4 trial:

```
metabias logor selogor if trial<16, graph(begg)
```

Note: default data input format (theta, se\_theta) assumed.

```
if trialno < 16
```

## META-ANALYSIS IN STATA

### Tests for Publication Bias

#### Begg's Test

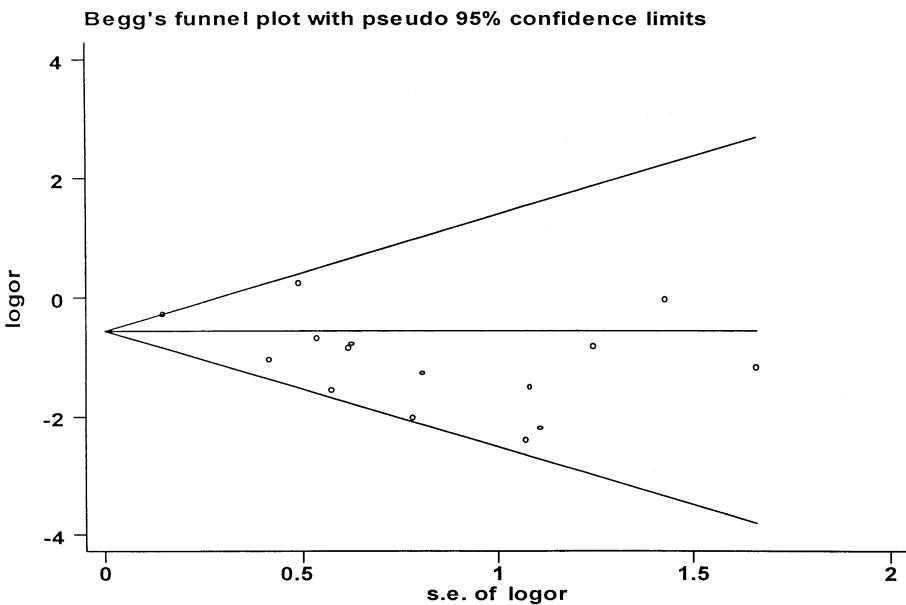
```

adj. Kendall's Score (P-Q) =      -3
    Std. Dev. of Score =   20.21
    Number of Studies =     15
        z =   -0.15
    Pr > |z| =   0.882
        z =   0.10 (continuity corrected)
    Pr > |z| =   0.921 (continuity corrected)

```

#### Egger's test

Std_Eff	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
slope	-.1512257	.1674604	-0.903	0.383	-.5130019 .2105505
bias	-1.192429	.3751749	-3.178	0.007	-2.002945 -.3819131



The funnel plot appears asymmetric, and there is evidence of bias using the Egger (weighted regression) method ( $P$  for bias 0.007) but not using the Begg (rank correlation method). This is compatible with a greater statistical power of the regression test, as discussed in Chapter 11. The horizontal line in the funnel plot indicates the fixed-effects summary estimate (using

The funnel plot appears asymmetric, and there is evidence of bias using the Egger (weighted regression) method ( $P$  for bias 0·007) but not using the Begg (rank correlation method). This is compatible with a greater statistical power of the regression test, as discussed in Chapter 11. The horizontal line in the funnel plot indicates the fixed-effects summary estimate (using inverse-variance weighting), while the sloping lines indicate the expected 95% confidence intervals for a given standard error, assuming no heterogeneity between studies.

### Meta-regression

If evidence is found of heterogeneity in the effect of treatment between studies, then meta-regression can be used to analyse associations between treatment effect and study characteristics. Meta-regression can be done in Stata by using the **metareg** command.<sup>19</sup>

#### Example 3: trials of BCG vaccine against tuberculosis

The following table is based on a meta-analysis by Colditz *et al.*<sup>20</sup> which examined the efficacy of BCG vaccine against tuberculosis.

Table 18.3

Trial	Trial name	Authors	Start year	Latitude*	Intervention group		Control group	
					TB cases	Total	TB cases	Total
1	Canada	Ferguson & Simes	1933	55	6	306	29	303
2	Northern USA	Aronson	1935	52	4	123	11	139
3	Northern USA	Stein & Aronson	1935	52	180	1541	372	1451
4	Chicago	Rosenthal <i>et al.</i>	1937	42	17	1716	65	1665
5	Chicago	Rosenthal <i>et al.</i>	1941	42	3	231	11	220
6	Georgia (School)	Comstock & Webster	1947	33	5	2498	3	2341
7	Puerto Rico	Comstock <i>et al.</i>	1949	18	186	50 634	141	27 338
8	UK	Hart & Sutherland	1950	53	62	13 598	248	12 867
9	Madanapalle	Frimont-Moller <i>et al.</i>	1950	13	33	5069	47	5808
10	Georgia (Community)	Comstock <i>et al.</i>	1950	33	27	16 913	29	17 854
11	Haiti	Vandeviere <i>et al.</i>	1965	18	8	2545	10	629
12	South Africa	Coetzee & Berjak	1965	27	29	7499	45	7277
13	Madras	TB prevention trial	1968	13	505	88 391	499	88 391

\* Expressed in degrees from equator.

The data were saved in Stata dataset **bcgtrial.dta**.

**describe**

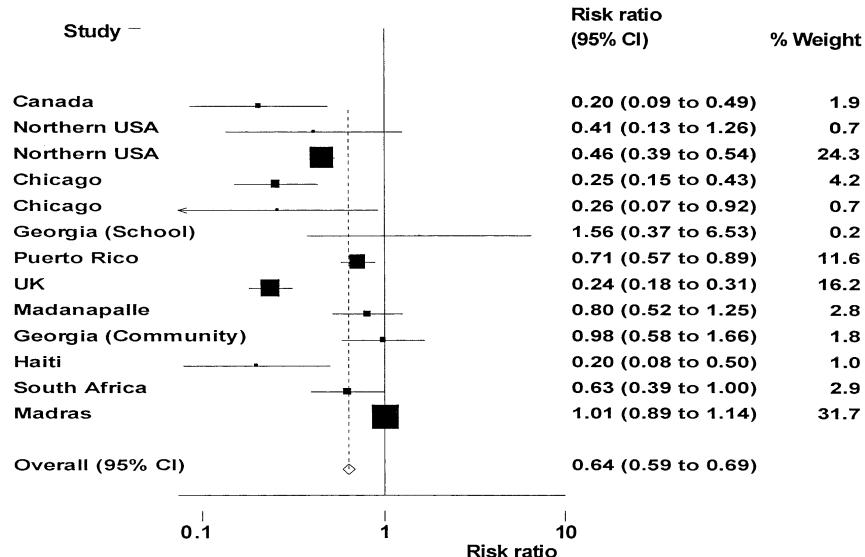
```
Contains data from bcgtrial.dta
obs:          13
vars:          9
size:        754 (99.9% of memory free)
-----
1. trial      byte    %8.0g
2. trialnam   str19   %19s
3. authors    str19   %19s
4. startyr    int     %8.0g
5. latitude   byte    %8.0g
6. cases1     int     %8.0g
7. tot1       long    %12.0g
8. cases0     int     %8.0g
9. tot0       long    %12.0g
-----
Sorted by: trial
```

Scientists had been aware of discordance between the results of these trials since the 1950s. The clear heterogeneity in the protective effect of BCG between trials can be seen in the forest plot (we analyse this study using risk ratios):

```
gen h1=tot1-cases1
gen h0=tot0-cases0

metan cases1 h1 cases0 h0, xlabel(.1,1,10)
label(namevar=trialnam)
```

## SYSTEMATIC REVIEWS IN HEALTH CARE



To use the **metareg** command, we need to derive the treatment effect estimate (in this case log risk ratio) and its standard error, for each study.

```
generate logrr=log((cases1/tot1)/(cases0/tot0))
generate selogrr=sqrt((1/cases1)-(1/tot1)+(1/cases0)-
(1/tot0))
```

In their meta-analysis, Colditz *et al.* noted the strong evidence for heterogeneity between studies, and concluded that a random-effects meta-analysis was appropriate:

```
meta logrr selogrr, eform
Meta-analysis (exponential form)
```

Method	Pooled		95% CI		Asymptotic		No. of studies
	Est	Lower	Upper	z_value	p_value		
Fixed	0.650	0.601	0.704	-10.625	0.000		13
Random	0.490	0.345	0.695	-3.995	0.000		

Test for heterogeneity: Q= 152.233 on 12 degrees of freedom (p= 0.000)

Moment-based estimate of between studies variance = 0.309

META-ANALYSIS IN STATA

(The different weight of studies under the fixed and random effects assumption is discussed in Chapter 2).

The authors then examined possible explanations for the clear differences in the effect of BCG between studies. The earlier studies may have produced different results than later ones. The latitude at which the studies were conducted may also be associated with the effect of BCG. As discussed by Fine,<sup>21</sup> the possibility that BCG might provide greater protection at higher latitudes was first recognised by Palmer and Long,<sup>22</sup> who suggested that this trend might result from exposure to certain environmental mycobacteria, more common in warmer regions, which impart protection against tuberculosis.

To use **metareg**, we provide a list of variables, the first of which is the treatment effect (here, the log risk ratio) and the rest of which are (one or more) study characteristics (covariates) hypothesized to be associated with the treatment effect. In addition, the standard error or variance of the treatment effect must be provided, using the **wsse** (within-study standard error) or **wsvar** (within-study variance) option. It is also possible to specify the method for estimating the between-study variance: here we use the default; restricted maximum-likelihood (reml). To look for an association with start year and latitude:

```
metareg logrr startyr latitude, wsse(selogrr)
```

```
Iteration 1: tau^2 = 0
Iteration 2: tau^2 = .02189942
:
:
Iteration 9: tau^2 = .1361904
Iteration 10: tau^2 = .13635174
```

Meta-analysis regression  
No of studies = 13  
 $\tau^2$  method reml  
 $\tau^2$  estimate = .1364

Successive values of tau^2 differ by less than 10^-4 :convergence achieved

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]
startyr	-.004966	.0162811	-0.305	0.760	-.0368763 .0269444
latitude	-.0270477	.0118195	-2.288	0.022	-.0502135 -.0038819
_cons	9.890987	32.02516	0.309	0.757	-52.87717 72.65914

#### SYSTEMATIC REVIEWS IN HEALTH CARE

The regression coefficients are the estimated increase in the log risk ratio per unit increase in the covariate. So in the example the log risk ratio is estimated to decrease by 0·027 per unit increase in the latitude at which the study is conducted. The estimated between-study variance has been reduced from 0·31 (see output from the **meta** command) to 0·14. While there is strong evidence for an association between latitude and the effect of BCG, there is no evidence for an association with the year the study started. The estimated treatment effect given particular values of the covariates may be derived from the regression equation. For example, for a trial beginning in 1950, at latitude 50°, the estimated log risk ratio is given by:

$$\text{Log risk ratio} = 9.891 - 0.00497 \times 1950 - 0.0270 \times 50 = -1.1505$$

which corresponds to a risk ratio of  $\exp(-1.1505) = 0.316$

The use of meta-regression in explaining heterogeneity and identifying sources of bias in meta-analysis is discussed further in Chapters 8–11.

- 1 Yusuf S, Collins R, Peto R, *et al*. Intravenous and intracoronary fibrinolytic therapy in acute myocardial infarction: overview of results on mortality, reinfarction and side-effects from 33 randomized controlled trials. *Eur Heart J* 1985;6:556–85.
- 2 Gruppo Italiano per lo Studio della Streptochinasi nell'Infarto Miocardico (GISSI). Effectiveness of intravenous thrombolytic treatment in acute myocardial infarction. *Lancet* 1986;1:397–402.
- 3 ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. *Lancet* 1988;2:349–60.
- 4 Bradburn MJ, Deeks JJ, Altman DG. sbe24: metan – an alternative meta-analysis command. *Stata Tech Bull* 1998;44:15.
- 5 Sharp S, Sterne J. sbe16: Meta-analysis. *Stata Tech Bull* 1997;38:9–14.
- 6 Sharp S, Sterne J. sbe16.1: New syntax and output for the meta-analysis command. *Stata Tech Bull* 1998;42:6–8.
- 7 Sharp S, Sterne J. sbe16.2: Corrections to the meta-analysis command. *Stata Tech Bull* 1998;43:15.
- 8 Teo KK, Yusuf S, Collins R, Held PH, Peto R. Effects of intravenous magnesium in suspected acute myocardial infarction: overview of randomised trials. *BMJ* 1991;303:1499–503.
- 9 ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 1995;345:669–85.
- 10 Egger M, Smith GD. Misleading meta-analysis. Lessons from an “effective, safe, simple” intervention that wasn’t. *BMJ* 1995;310:752–4.
- 11 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
- 12 Sterne J. sbe22: Cumulative meta analysis. *Stata Tech Bull* 1998;42:13–16.
- 13 Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992;327:248–54.
- 14 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts’ Treatments for myocardial infarction. *JAMA* 1992;268:240–8.
- 15 Tobias A. sbe26: Assessing the influence of a single study in meta-analysis. *Stata Tech Bull* 1999;47:15–17.

#### META-ANALYSIS IN STATA

- 16 Steichen T. sbe19: Tests for publication bias in meta-analysis. *Stata Tech Bull* 1998;41:9–15.
- 17 Steichen T, Egger M, Sterne J. sbe19.1: Tests for publication bias in meta-analysis. *Stata Tech Bull* 1998;44:3–4.
- 18 Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics* 1994;50:1088–101.
- 19 Sharp S. sbe23: Meta-analysis regression. *Stata Tech Bull* 1998;42:16–24.
- 20 Colditz GA, Brewer TF, Berkey CS, et al. Efficacy of BCG vaccine in the prevention of tuberculosis. Meta-analysis of the published literature. *JAMA* 1994;271:698–702.
- 21 Fine PEM. Variation in protection by BCG: implications of and for heterologous immunity. *Lancet* 1995;346:1339–45.
- 22 Palmer CE, Long MW. Effects of infection with atypical mycobacteria on BCG vaccination and tuberculosis. *Am Rev Respir Dis* 1966;94:553–68.