metan: fixed- and random-effects meta-analysis

Ross J. Harris
Department of Social Medicine
University of Bristol
Bristol, UK
epzrgh@bristol.ac.uk

Michael J. Bradburn
Health Services Research Center
University of Sheffeld
Sheffield, UK

Jonathan J. Deeks
Department of Primary Care Medicine
University of Birmingham
Birmingham, UK

Roger M. Harbord
Department of Social Medicine
University of Bristol
Bristol, UK

Douglas G. Altman
Centre for Statistics in Medicine
University of Oxford
Oxford, UK

Jonathan A. C. Sterne
Department of Social Medicine
University of Bristol
Bristol, UK

Abstract. This article describes updates of the meta-analysis command `metan` and options that have been added since the command's original publication (Bradburn, Deeks, and Altman, metan – an alternative meta-analysis command, *Stata Technical Bulletin Reprints*, vol. 8, pp. 86–100). These include version 9 graphics with flexible display options, the ability to meta-analyze precalculated effect estimates, and the ability to analyze subgroups by using the `by()` option. Changes to the output, saved variables, and saved results are also described.

Keywords: sbe24_2, metan, meta-analysis, forest plot

1 Introduction

Meta-analysis is a two-stage process involving the estimation of an appropriate summary statistic for each of a set of studies followed by the calculation of a weighted average of these statistics across the studies (Deeks, Altman, and Bradburn 2001). Odds ratios, risk ratios, and risk differences may be calculated from binary data, or a difference in means obtained from continuous data. Alternatively, precalculated effect estimates and their standard errors from each study may be pooled, for example, adjusted log-odds ratios from observational studies. The summary statistics from each study can be combined by using a variety of meta-analytic methods, which are classified as fixed-effect models in which studies are weighted according to the amount of information they contain; or random-effects models, which incorporate an estimate of between-study variation (heterogeneity) in the weighting. A meta-analysis will customarily include a forest plot, in which results from each study are displayed as a square and a horizontal line, representing the intervention effect estimate together with its confidence interval. The area of the square reflects the weight that the study contributes to the meta-

analysis. The combined-effect estimate and its confidence interval are represented by a diamond.

Here we present updates to the `metan` command and other previously undocumented additions that have been made since its original publication (Bradburn, Deeks, and Altman 1998). New features include

- Version 9 graphics

- Flexible display of tabular data in the forest plot

- Results from a second type of meta-analysis displayed in the same forest plot

- `by()` group processing

- Analysis of precalculated effect estimates

- Prediction intervals for the intervention effect in a new study from random-effects analyses

There are a substantial number of options for the `metan` command because of the variety of meta-analytic techniques and the need for flexible graphical displays. We recommend that new users not try to learn everything at once but to learn the basics and build from there as required. Clickable examples of `metan` are available in the help file, and the dialog box may also be a good way to start using `metan`.

2   Example data

The dataset used in subsequent examples is taken from the meta-analysis published as table 1 in Colditz et al. (1994, 699). The aim of the analysis was to quantify the efficacy of BCG vaccine against tuberculosis, and data from 11 trials are included here. There was considerable between-trial heterogeneity in the effect of the vaccine; it has been suggested that this might be explained by the latitude of the region in which the trial was conducted (Fine 1995).

▷ **Example**

Details of the dataset are shown below by using `describe` and `list` commands.

```
. use bcgtrial
(BCG and tuberculosis)

. describe

Contains data from bcgtrial.dta
  obs:            11                          BCG and tuberculosis
 vars:            12                          31 May 2007 17:11
 size:           693 (99.9% of memory free)  (_dta has notes)

              storage  display    value
variable name   type   format     label      variable label

trial          byte    %8.0g                 Trial number
trialnam       str14   %14s                  Trial name
authors        str20   %20s                  Authors of trial
startyr        int     %8.0g                 Year trial started
latitude       byte    %8.0g                 Latitude of trial area
alloc          byte    %33.0g     alloc      Allocation method
tcases         int     %8.0g                 BCG vaccinated cases
tnoncases      float   %9.0g                 BCG vaccinated noncases
ccases         int     %8.0g                 Unvaccinated cases
cnoncases      float   %9.0g                 Unvaccinated noncases
ttotal         long    %12.0g                BCG vaccinated population
ctotal         long    %12.0g                Unvaccinated population

Sorted by:  startyr  authors

. list trialnam startyr tcases tnoncases ccases cnoncases, clean noobs
> abbreviate(10)
            trialnam  startyr  tcases  tnoncases  ccases  cnoncases
              Canada     1933       6        300      29        274
       Northern USA     1935       4        119      11        128
             Chicago     1941      17       1699      65       1600
       Georgia (Sch)     1947       5       2493       3       2338
         Puerto Rico     1949     186      50448     141      27197
      Georgia (Comm)     1950      27      16886      29      17825
         Madanapalle     1950      33       5036      47       5761
                  UK     1950      62      13536     248      12619
        South Africa     1965      29       7470      45       7232
               Haiti     1965       8       2537      10        619
              Madras     1968     505      87886     499      87892
```

Trial name and number identify each study, and we have information on the authors and the year the trial started. There are also two variables relating to study characteristics: the latitude of the area in which the trial was carried out, and the method of allocating patients to the vaccine and control groups—either at random or in some systematic way. The variables `tcases`, `tnoncases`, `ccases`, and `cnoncases` contain the data from the 2×2 table from each study (the number of cases and noncases in the vaccination group and nonvaccination group). The variables `ttotal` and `ctotal` are the total number of individuals (the sum of the cases and noncases) in the vaccine and control groups. Displayed below is the 2×2 table for the first study (Canada, 1933):

	cases	noncases	total
treated	6	300	306
control	29	274	303

The risk ratio (RR), log-risk ratio (log-RR), standard error of log-RR (SE log-RR), 95% confidence interval (CI) for log-RR, and 95% CI for RR may be calculated as follows (see, for example, Kirkwood and Sterne 2003).

$$\text{Risk in treated population} = \frac{\texttt{tcases}}{\texttt{ttotal}} = \frac{6}{306} = 0.0196$$

$$\text{Risk in control population} = \frac{\texttt{ccases}}{\texttt{ctotal}} = \frac{29}{303} = 0.0957$$

$$\text{RR} = \frac{\text{Risk in treated population}}{\text{Risk in control population}} = \frac{0.0196}{0.0957} = 0.2049$$

$$\log \text{RR} = \log(\text{RR}) = -1.585$$

$$\text{SE}(\log \text{RR}) = \sqrt{\frac{1}{\texttt{tcases}} + \frac{1}{\texttt{ccases}} - \frac{1}{\texttt{ttotal}} - \frac{1}{\texttt{ctotal}}}$$

$$= \sqrt{\frac{1}{6} + \frac{1}{29} - \frac{1}{306} - \frac{1}{303}} = 0.441$$

$$95\% \text{ CI for log RR} = \log \text{RR} \pm 1.96 \times \text{SE}(\log \text{RR}) = -2.450 \text{ to } -0.720$$

$$95\% \text{ CI for RR} = \exp(-2.450) \text{ to } \exp(-0.720) = 0.086 \text{ to } 0.486$$

◁

3   Syntax

`metan` *varlist* $\big[\,$*if*$\,\big]$ $\big[\,$*in*$\,\big]$ $\big[\,$,

　　$\big[\,$*binary_data_options* | *continuous_data_options* | *precalculated_effect_estimates_options* $\big]$

　　measure_and_model_options output_options forest_plot_options $\big]$

binary_data_options

　　or rr rd fixed random fixedi randomi peto cornfield chi2 breslow
　　<u>noint</u>eger cc(#)

continuous_data_options

　　cohen hedges glass nostandard fixed random <u>noint</u>eger

precalculated_effect_estimates_options

> fixed random

measure_and_model_options

> wgt(*wgtvar*) second(*model | estimates_and_description*)
> first(*estimates_and_description*)

output_options

> by(*byvar*) nosubgroup sgweight log eform efficacy ilevel(*#*)
> olevel(*#*) sortby(*varlist*)
> label([namevar = *namevar*], [yearvar = *yearvar*]) nokeep notable nograph
> nosecsub

forest_plot_options

> xlabel(*#*, ...) xtick(*#*, ...) boxsca(*#*) textsize(*#*) nobox nooverall
> nowt nostats counts group1(*string*) group2(*string*) effect(*string*) force
> lcols(*varlist*) rcols(*varlist*) astext(*#*) double nohet summaryonly rfdist
> rflevel(*#*) null(*#*) nulloff favours(*string # string*) firststats(*string*)
> secondstats(*string*) boxopt(*marker_options*) diamopt(*line_options*)
> pointopt(*marker_options | marker_label_options*) ciopt(*line_options*)
> olineopt(*line_options*) classic nowarning *graph_options*

For a full description of the syntax, see Bradburn, Deeks, and Altman (1998). We will focus on the new options, most of which come under *forest_plot_options*; previously undocumented options such as by() (and related options), breslow, cc(), nointeger; and changes to the output such as the display of the I^2 statistic. Syntax will be explained in the appropriate sections.

4   Basic use

4.1 2×2 data

For binary data, the input variables required by metan should contain the cells of the 2×2 table; i.e., the number of individuals who did and did not experience the outcome event in the treatment and control groups for each study. When analyzing 2×2 data a range of methods are available. The default is the Mantel–Haenszel method (fixed). The inverse-variance fixed-effect method (fixedi) or the Peto method for estimating summary odds ratios (peto) may also be chosen. The DerSimonian and Laird random-effects method may be specified with random. See Deeks, Altman, and Bradburn (2001) for a discussion of these methods.

4.2   Display options

Previous versions of the `metan` command used the syntax `label(namevar = ` *namevar* `,` `yearvar = ` *yearvar*`)` to specify study information in the table and forest plot. This syntax still functions but has been superseded by the more flexible `lcols(`*varlist*`)` and `rcols(`*varlist*`)` options. The use of these options is described in more detail in section 5. The option `favours(`*string* `#` *string*`)` allows the user to display text information about the direction of the treatment effect, which appears under the graph (e.g., exposure good, exposure bad). `favours()` replaces the option `b2title()`. The `#` is required to split the two strings, which appear to either side of the null line.

▷ **Example**

Here we use `metan` to derive an inverse-variance weighted (fixed effect) meta-analysis of the BCG trial data. Risk ratios are specified as the summary statistic, and the trial name and the year the trial started are displayed in the forest plot using `lcols()` (see section 5).

```
. metan tcases tnoncases ccases cnoncases, rr fixedi lcols(trialnam startyr)
> xlabel(0.1, 10) favours(BCG reduces risk of TB # BCG increases risk of TB)
```

Study	RR	[95% Conf. Interval]		% Weight
Canada	0.205	0.086	0.486	1.11
Northern USA	0.411	0.134	1.257	0.66
Chicago	0.254	0.149	0.431	2.96
Georgia (Sch)	1.562	0.374	6.528	0.41
Puerto Rico	0.712	0.573	0.886	17.42
Georgia (Comm)	0.983	0.582	1.659	3.03
Madanapalle	0.804	0.516	1.254	4.22
UK	0.237	0.179	0.312	10.81
South Africa	0.625	0.393	0.996	3.83
Haiti	0.198	0.078	0.499	0.97
Madras	1.012	0.895	1.145	54.58
I-V pooled RR	0.730	0.667	0.800	100.00

```
  Heterogeneity chi-squared = 125.63 (d.f. = 10) p = 0.000
  I-squared (variation in RR attributable to heterogeneity) =  92.0%
  Test of RR=1 : z=   6.75 p = 0.000
```

The output table contains effect estimates (here, RRs), CIs, and weights for each study, followed by the overall (combined) effect estimate. The results for the Canada study are identical to those derived in section 2. Heterogeneity statistics relating to the extent that RRs vary between studies are displayed, including the I^2 statistic, which is a previously undocumented addition. The I^2 statistic (see section 9.1) is the percentage of between-study heterogeneity that is attributable to variability in the true treatment effect, rather than sampling variation (Higgins and Thompson 2004, Higgins et al. 2003). Here there is substantial between-study heterogeneity. Finally, a test of the null hypothesis that the vaccine has no effect (RR=1) is displayed. There is strong evidence against the null hypothesis, but the presence of between-study heterogeneity means that

the fixed-effect assumption (that the true treatment effect is the same in each study) is incorrect. The forest plot displayed by the command is shown in figure 1.
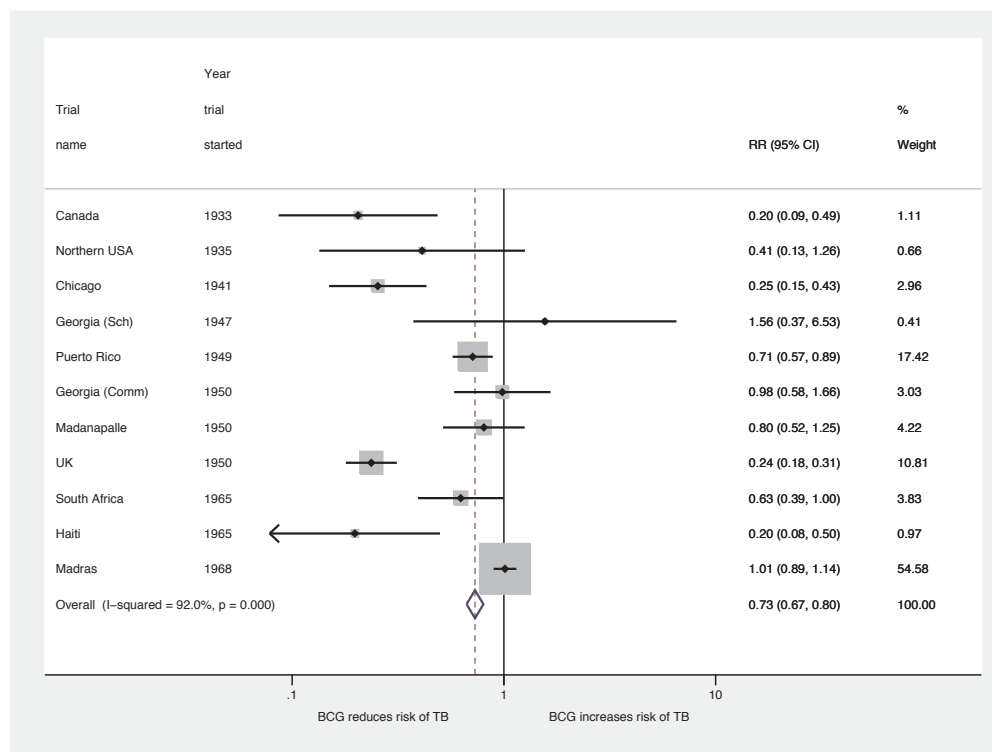

Figure 1. Forest plot displaying an inverse-variance weighted fixed-effect meta-analysis of the effect of BCG vaccine on incidence of tuberculosis.

◁

4.3 Precalculated effect estimates

The `metan` command may also be used to meta-analyze precalculated effect estimates, such as log-odds ratios and their standard errors or 95% CI, using syntax similar to the alternative Stata meta-analysis command `meta` (Sharp and Sterne 1997). Here only the inverse-variance fixed-effect and DerSimonian and Laird random-effects methods are available, because other methods require the 2×2 cell counts or the means and standard deviations in each group. The `fixed` option produces an inverse-variance weighted analysis when precalculated effect estimates are analyzed.

When analyzing ratio measures (RRs or odds ratios), the log ratio with its standard error or 95% CI should be used as inputs to the command. The `eform` option can then be used to display the output on the ratio scale (as for the `meta` command).

▷ **Example**

We will illustrate this feature by generating the log-RR and its standard error in each study from the 2×2 data, and then by meta-analyzing these variables.

```
. gen logRR = ln( (tcases/ttotal) / (ccases/ctotal) )
. gen selogRR = sqrt( 1/tcases +1/ccases -1/ttotal -1/ctotal )
. metan logRR selogRR, fixed eform nograph
```

Study	ES	[95% Conf. Interval]	% Weight
(table of study results omitted)			
I-V pooled ES	0.730	0.667 0.800	100.00

```
  Heterogeneity chi-squared = 125.63 (d.f. = 10) p = 0.000
  I-squared (variation in ES attributable to heterogeneity) =  92.0%
  Test of ES=1 : z=   6.75 p = 0.000
```

The results are identical to those derived directly from the 2×2 data in section 4.1; we would have observed minor differences if the default Mantel–Haenszel method had been used previously. When analyzing precalculated estimates, metan does not know what these measures are, so the summary estimate is named "ES" (effect size) in the output.

◁

4.4   Specifying two analyses

metan now allows the display of a second meta-analytic estimate in the same output table and forest plot. A typical use is to compare fixed-effect and random-effects analyses, which can reveal the presence of small-study effects. These may result from publication or other biases (Sterne, Gavaghan, and Egger 2000). See Poole and Greenland (1999) for a discussion of the ways in which fixed-effect and random-effects analyses may differ. The syntax is to specify the method for the second meta-analytic estimate as second(*method*), where *method* is any of the standard metan options.

▷ **Example**

Here we use metan to analyze 2×2 data as in section 4.1, specifying an inverse-variance weighted (fixed effect) model for the first method and a DerSimonian and Laird (random effects) model for the second method:

```
. metan tcases tnoncases ccases cnoncases, rr fixedi second(random)
> lcols(trialnam startyr) nograph
         Study       |   RR    [95% Conf. Interval]    % Weight
```

                     (table of study results omitted)

```
I-V pooled RR        |  0.730    0.667    0.800        100.00
D+L pooled RR        |  0.508    0.336    0.769        100.00
```

```
  Heterogeneity chi-squared = 125.63 (d.f. = 10) p = 0.000
  I-squared (variation in RR attributable to heterogeneity) =  92.0%

  Test of RR=1 : z=   6.75 p = 0.000
```

The results of the second analysis are displayed in the table: a forest plot using the `second()` option is derived in the next section and displayed in figure 2. The protective effect of BCG against tuberculosis appears greater in the random-effects analysis than in the fixed-effect analysis, although CI is wider. This reflects the greater uncertainty in the random-effects analysis, which allows for the true effect of the vaccine to vary between studies. Random-effects analyses give relatively greater weight to smaller studies than fixed-effect analyses, and so these results suggest that the estimated effect of BCG was greater in the smaller studies. It is also possible to supply a precalculated pooled-effect estimate with `second()`; see section 7.2 for details.

◁

5   Displaying data columns in graphs

The options `lcols(`*varlist*`)` and `rcols(`*varlist*`)` produce columns to the left or right of the forest plot. String (character) or numeric variables can be displayed. If numeric variables have value labels, these will be displayed in the graph. If the variable itself is labeled, this will be used as the column header, allowing meaningful names to be used. Up to four lines are used for the heading, so names can be long without taking up too much graph width.

The first variable in `lcols()` is used to identify studies in the table output, and summary statistics and study weight are always the first columns on the right of the forest plot. These can be switched off by using the options `nostats` and `nowt`, but the order cannot be changed.

If lengthy string variables are to be displayed, the `double` option may be used to allow output to spread over two lines per study in the forest plot. The percentage of the forest plot given to text may be adjusted using `astext(`#`)`, which can be between 10 and 90 (the default is 50).

A previously undocumented option that affects columns is `counts`. When this option is specified, more columns will appear on the right of the graph displaying the raw data; either the 2×2 table for binary data or the sample size, mean, and standard deviation in each group if the data are continuous. The groups may be labeled by using `group1(`*string*`)` and `group2(`*string*`)`, although the defaults *Treatment* and *Control* will often be acceptable for the analysis of randomized controlled trials (RCTs).

▷ **Example**

We now present an example command that uses these features, as well as the
`second()` option. The resulting forest plot is displayed in figure 2:

```
. metan tcases tnoncases ccases cnoncases, rr fixedi second(random)
> lcols(trialnam authors startyr alloc latitude) counts astext(70)
> textsize(200) boxsca(80) xlabel(0.1,10) notable xsize(10) ysize(6)
```


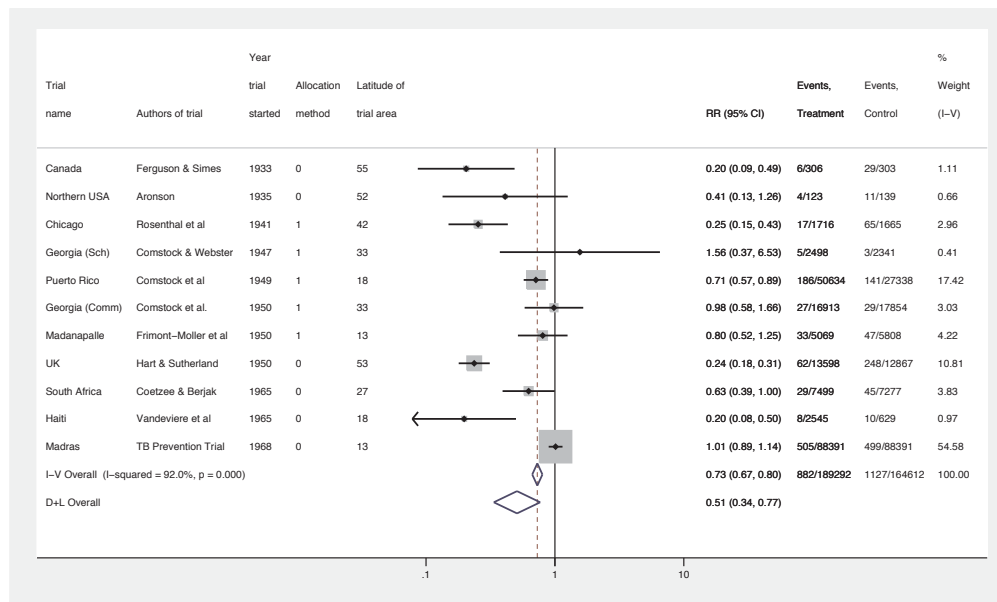| Trial name	Authors of trial	Year trial started	Allocation method	Latitude of trial area		RR (95% CI)	Events, Treatment	Events, Control	% Weight (I–V)
Canada	Ferguson & Simes	1933	0	55		0.20 (0.09, 0.49)	6/306	29/303	1.11
Northern USA	Aronson	1935	0	52		0.41 (0.13, 1.26)	4/123	11/139	0.66
Chicago	Rosenthal et al	1941	1	42		0.25 (0.15, 0.43)	17/1716	65/1665	2.96
Georgia (Sch)	Comstock & Webster	1947	1	33		1.56 (0.37, 6.53)	5/2498	3/2341	0.41
Puerto Rico	Comstock et al	1949	1	18		0.71 (0.57, 0.89)	186/50634	141/27338	17.42
Georgia (Comm)	Comstock et al.	1950	1	33		0.98 (0.58, 1.66)	27/16913	29/17854	3.03
Madanapalle	Frimont–Moller et al	1950	1	13		0.80 (0.52, 1.25)	33/5069	47/5808	4.22
UK	Hart & Sutherland	1950	0	53		0.24 (0.18, 0.31)	62/13598	248/12867	10.81
South Africa	Coetzee & Berjak	1965	0	27		0.63 (0.39, 1.00)	29/7499	45/7277	3.83
Haiti	Vandeviere et al	1965	0	18		0.20 (0.08, 0.50)	8/2545	10/629	0.97
Madras	TB Prevention Trial	1968	0	13		1.01 (0.89, 1.14)	505/88391	499/88391	54.58
I–V Overall (I–squared = 92.0%, p = 0.000)						0.73 (0.67, 0.80)	882/189292	1127/164612	100.00
D+L Overall						0.51 (0.34, 0.77)			

Figure 2. Forest plot displaying an inverse-variance weighted fixed-effect meta-analysis
of the effect of BCG vaccine on incidence of tuberculosis. Columns of data are displayed
in the plot.

Note the specification of x-axis labels and text and box sizes. The graph is also reshaped
by using the standard Stata graph options `xsize()` and `ysize()`; see section 10.2 for
more details. Box and text sizes are expressed as a percentage of standard size with the
default as 100, such that 50 will halve the size and 200 will double it.

◁

6   by() processing

A major addition to `metan` is the ability to perform stratified or subgroup analyses.
These may be used to investigate the possibility that treatment effects vary between
subgroups; however, formal comparisons between subgroups are best performed by using
meta-regression; see Harbord and Higgins (2008) or Higgins and Thompson (2004). We

may also want to display results for different groups of studies in the same plot, even though it is inappropriate to meta-analyze across these groups.

6.1   Syntax and options for by()

nooverall specifies that the overall estimate not be displayed, for example, when it is inappropriate to meta-analyze across groups.

sgweight requests that weights be displayed such that they sum to 100% within each subgroup. This option is invoked automatically with **nooverall**.

nosubgroup specifies that studies be arranged by the subgroup specified, but estimates for each subgroup not be displayed.

nosecsub specifies that subestimates using the method defined by **second()** not be displayed.

summaryonly specifies that individual study estimates not be displayed, for example, to produce a summary of different groups in a compact graph.

▷ **Example**

Fine (1995) suggested that there is a relationship between the effect of BCG and the latitude of the area in which the trial was conducted. Here we may want to use meta-regression to further investigate this tendency (see Harbord and Higgins 2008). To illustrate the by() option, we will classify the studies into three groups defined by latitude. We define these groups as tropical ($\leq$23.5 degrees), midlatitude (between 23.5 and 40 degrees) and northern ($\geq$40 degrees).

```
. gen lat_cat = ""
(11 missing values generated)
. replace lat_cat = "Tropical, < 23.5 latitude" if latitude <= 23.5
lat_cat was str1 now str27
(4 real changes made)
. replace lat_cat = "23.5-40 latitude" if latitude > 23.5 & latitude < 40
(3 real changes made)
. replace lat_cat = "Northern, > 40 latitude" if latitude >= 40 & latitude < .
(4 real changes made)
. assert lat_cat != ""
. label var lat_cat "Latitude region"
```

(Continued on next page)

```
. metan tcases tnoncases ccases cnoncases, rr fixedi second(random) nosecsub
> lcols(trialnam startyr latitude) astext(60) by(lat_cat) xlabel(0.1,10)
> xsize(10) ysize(8)
```

Study	RR	[95% Conf. Interval]		% Weight
Northern, > 40 lat				
Canada	0.205	0.086	0.486	1.11
Northern USA	0.411	0.134	1.257	0.66
Chicago	0.254	0.149	0.431	2.96
UK	0.237	0.179	0.312	10.81
Sub-total				
I-V pooled RR	0.243	0.193	0.306	15.54
23.5-40 latitude				
Georgia (Sch)	1.562	0.374	6.528	0.41
Georgia (Comm)	0.983	0.582	1.659	3.03
South Africa	0.625	0.393	0.996	3.83
Sub-total				
I-V pooled RR	0.795	0.567	1.114	7.27
Tropical, < 23.5 l				
Puerto Rico	0.712	0.573	0.886	17.42
Madanapalle	0.804	0.516	1.254	4.22
Haiti	0.198	0.078	0.499	0.97
Madras	1.012	0.895	1.145	54.58
Sub-total				
I-V pooled RR	0.904	0.815	1.003	77.19
Overall				
I-V pooled RR	0.730	0.667	0.800	100.00
D+L pooled RR	0.508	0.336	0.769	

```
Test(s) of heterogeneity:
                  Heterogeneity   degrees of
                    statistic      freedom        P     I-squared**
Northern, > 40 lat   1.06             3         0.787      0.0%
23.5-40 latitude     2.51             2         0.285     20.2%
Tropical, < 23.5 l  18.42             3         0.000     83.7%
Overall            125.63            10         0.000     92.0%
Overall Test for heterogeneity between sub-groups:
                   103.64             2         0.000

** I-squared: the variation in RR attributable to heterogeneity)

Considerable heterogeneity observed (up to 83.7%) in one or more sub-groups,
Test for heterogeneity between sub-groups likely to be invalid

Significance test(s) of RR=1

Northern, > 40 lat  z= 12.00     p = 0.000
23.5-40 latitude    z=  1.33     p = 0.183
Tropical, < 23.5 l  z=  1.90     p = 0.058
Overall             z=  6.75     p = 0.000
```
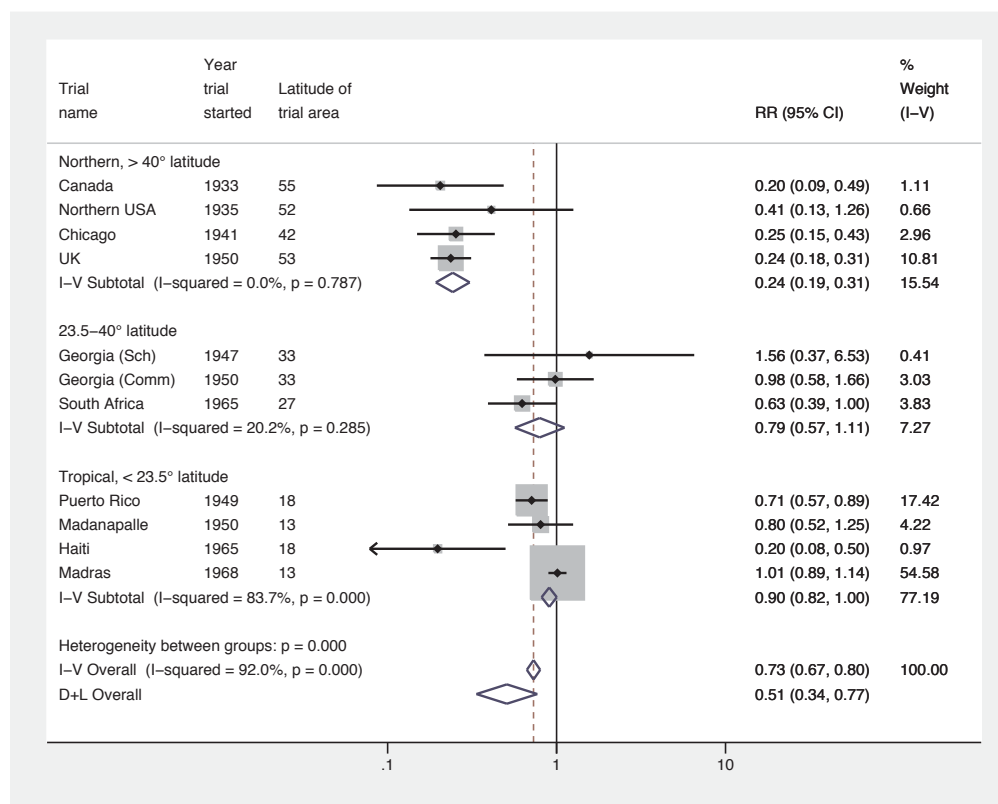
Figure 3. Forest plot displaying an inverse-variance weighted fixed-effect meta-analysis of the effect of BCG vaccine on incidence of tuberculosis. Results are stratified by latitude region, and the overall random-effects estimate is also displayed.

The output table is now stratified by latitude group, and pooled estimates for each group are displayed. Tests of heterogeneity and the null hypothesis are displayed for each group and overall. With the inverse-variance method, a test of heterogeneity between groups is also displayed; note the warning in the output that the test may be invalid because of within-subgroup heterogeneity. Output is similar in the forest plot, displayed in figure 3. Examining each subgroup in turn, it appears that much of the heterogeneity is accounted for by latitude: for two of the groups there is little or no evidence of heterogeneity. The only group to show a strong treatment effect is the ≥ 40 degree group.

◁

The test for between-group heterogeneity is an issue of current debate, as it is strictly valid only when using the fixed-effect inverse-variance method, and p-values will be too small if there is heterogeneity within any of the subgroups. Therefore, the test is performed only with the inverse-variance method (`fixedi`), and warnings will appear

if there is evidence of within-group heterogeneity. Despite these caveats, this method is better than other, seriously flawed, methods such as testing the significance of a treatment effect in each group rather than testing for differences between the groups. As explained at the start of this section, meta-regression is the best way to examine and test for between-group differences.

7   User-defined analyses

7.1   Study weights

The wgt(*wgtvar*) option allows the studies to be combined by using specific weights that are defined by the variable *wgtvar*. The user must ensure that the weights chosen are meaningful. Typical uses are when analyzing precalculated effect estimates that require weights that are not based on standard error or to assess the robustness of conclusions by assigning alternative weights.

7.2   Pooled estimates

Pooled estimates may be derived by using another package and presented in a forest plot by using the first() option to supply these to the metan command. Here wgt(*wgtvar*) is used merely to specify box sizes in the forest plot, no heterogeneity statistics are produced, and no values are returned. When using this feature, stratified analyses are not allowed.

An alternative method is to provide the user-supplied meta-analytic estimate by using the second() option. Data are analyzed by using standard methods, and the resulting pooled estimate is displayed together with the user-defined estimate (which need not be derived by using metan), allowing a comparison. When using this feature, the option nosecsub is invoked, as stratification using the user-defined method is not possible.

When these options are specified, the user must supply the pooled estimate with its standard error or CI and a method label. The user may also supply text to be displayed at the bottom of the forest plot, in the position normally given to heterogeneity statistics, using firststats(*string*) and secondstats(*string*).

▷ **Example**

The BCG data were analyzed by using a fully Bayesian random-effects model with WinBUGS software (Lunn et al. 2000). This analysis used the methods described by Warn, Thompson, and Spiegelhalter (2002) to deal with RRs. The chosen model incorporated a noninformative prior (mean 0, precision 0.001). The resulting RR of 0.518 (95% CI: 0.300, 0.824) is similar to that derived from a DerSimonian and Laird random-effects analysis. However, the CI from the Bayesian analysis is wider, because it allows for the uncertainty in estimating the between-study variance. The following syntax sup-

plies the summary estimates in `second()` and compares this result with the random-effects analysis. The resulting forest plot is displayed in figure 4.

```
. metan logRR selogRR, random second(-.6587 -1.205 -.1937 Bayes)
> secondstats(Noninformative prior: d~dnorm(0.0, 0.001)) eform
> notable astext(60) textsize(130) lcols(trialnam startyr latitude)
> xlabel(0.1,10)
```
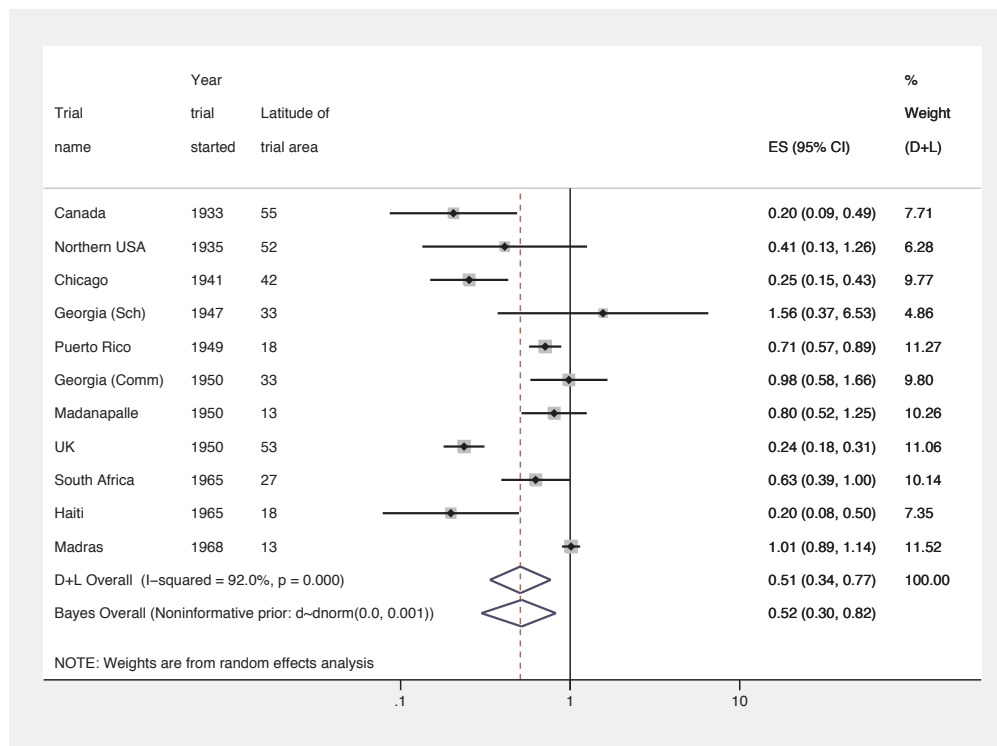

Figure 4. Forest plot displaying a fully Bayesian meta-analysis of the effect of BCG vaccine on incidence of tuberculosis. A noninformative prior has been specified, resulting in a pooled-effect estimate similar to the random-effects analysis.

◁

8   New analysis options

Here we discuss previously undocumented options added to `metan` since its original publication.

8.1   Dealing with zero cells

The cc(#) option allows the user to choose what value (if any) is to be added to the cells of the 2×2 table for a study in which one or more of the cell counts equals zero. Here the default is to add 0.5 to all cells of the 2×2 table for the study (except for the Peto method, which does not require a correction). This approach has been criticized, and other approaches (including making no correction) may be preferable (see Sweeting, Sutton, and Lambert [2004] for a discussion). The number declared in cc(#) must be between zero and one and will be added to each cell. When no events are recorded and RRs or odds ratios are to be combined the study is omitted, although for risk differences the effect is still calculable and the study is included. If no adjustment is made in the presence of zero cells, odds ratios and their standard errors cannot be calculated. Risk ratios and their standard errors cannot be calculated when the number of events in either the treatment or control group is zero.

8.2   Noninteger sample size

The nointeger option allows the number of observations in each arm (cell counts for binary data or the number of observations for continuous data) to be noninteger. By default, the sample size is assumed to be a whole number for both binary and continuous data. However, it may make sense for this not to be so, for example, to use a more flexible continuity correction with a different number added to each cell or when the meta-analysis incorporates cluster randomized trials and the effective-sample size is less than the total number of observations.

8.3   Breslow and Day test for heterogeneity

The breslow option can be used to perform the Breslow–Day test for heterogeneity of the odds ratio (Breslow and Day 1980). A review article by Reis, Hirji, and Afifi (1999) compared several different tests of heterogeneity and found this test to perform well in comparison to other asymptotic tests.

9   New output

9.1   The I^2 statistic

metan now displays the I^2 statistic as well as Cochran's Q to quantify heterogeneity, based on the work by Higgins and Thompson (2004) and Higgins et al. (2003). Briefly, I^2 is the percentage of variation attributable to heterogeneity and is easily interpretable. Cochran's Q can suffer from low power when the number of studies is low or excessive power when the number of studies is large. I^2 is calculated from the results of the meta-analysis by

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

where Q is Cochran's heterogeneity statistic and df is the degrees of freedom. Negative values of I^2 are set to zero so that I^2 lies between 0% and 100%. A value of 0% indicates no observed heterogeneity, and larger values show increasing heterogeneity. Although there can be no absolute rule for when heterogeneity becomes important, Higgins et al. (2003) tentatively suggest adjectives of low for I^2 values between 25%–50%, moderate for 50%–75%, and high for $\geq 75\%$.

9.2   Prediction interval for the random-effects distribution

The presentation of summary random-effects estimates may sometimes be misleading, as the CI refers to the average true treatment effect, but this is assumed under the random-effects model to vary between studies. A CI derived from a larger number of studies exhibiting a high degree of heterogeneity could be of similar width to a CI derived from a smaller number of more homogeneous studies, but in the first situation, we will be much less sure of the range within which the treatment effect in a new study will lie (Higgins and Thompson 2001). The prediction interval for the treatment effect in a new trial may be approximated by using the formula

$$\text{mean} \pm t_{\text{df}} \times \sqrt{(\text{se}^2 + \tau^2)}$$

where t is the appropriate centile point (e.g., 95%) of the t distribution with $k-2$ degrees of freedom, se^2 is the squared standard error, and τ^2 the between-study variance. This incorporates uncertainty in the location and spread of the random-effects distribution. The approximate prediction interval can be displayed in the forest plot, with lines extending from the summary diamond, by using the option `rfdist`. With ≤ 2 studies, the distribution is inestimable and effectively infinite; thus the interval is displayed with dotted lines. When heterogeneity is estimated to be zero, the prediction interval is still slightly wider than the summary diamond as the t statistic is always greater than the corresponding normal deviate. The coverage (e.g., 90%, 95%, or 99%) for the interval may be set by using the command `rflevel(#)`.

▷ **Example**

Here we display the prediction intervals corresponding to the stratified analyses derived in section 6.1. The resulting forest plot is displayed in figure 5.

```
. metan tcases tnoncases ccases cnoncases, rr random rfdist
> lcols(trialnam startyr latitude) astext(60) by(lat_cat) xlabel(0.1,10)
> xsize(10) ysize(8) notable
```

(*Continued on next page*)

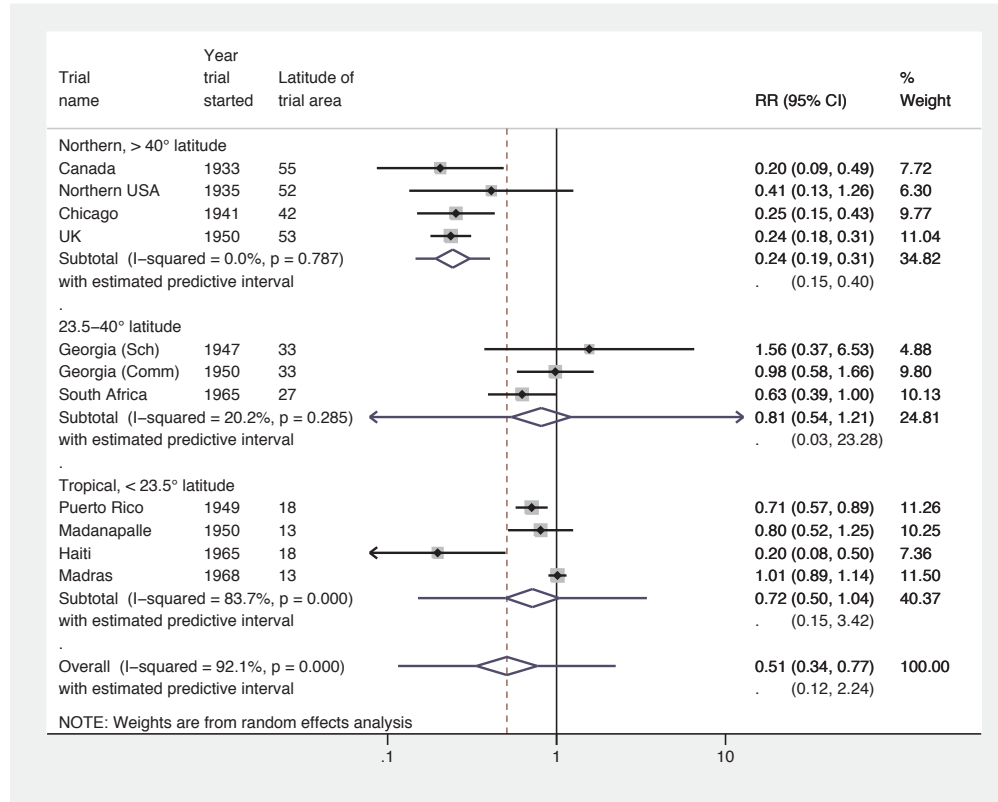| Trial name	Year trial started	Latitude of trial area		RR (95% CI)	% Weight
Northern, > 40° latitude					
Canada	1933	55		0.20 (0.09, 0.49)	7.72
Northern USA	1935	52		0.41 (0.13, 1.26)	6.30
Chicago	1941	42		0.25 (0.15, 0.43)	9.77
UK	1950	53		0.24 (0.18, 0.31)	11.04
Subtotal (I–squared = 0.0%, p = 0.787)				0.24 (0.19, 0.31)	34.82
with estimated predictive interval				. (0.15, 0.40)	
.					
23.5–40° latitude					
Georgia (Sch)	1947	33		1.56 (0.37, 6.53)	4.88
Georgia (Comm)	1950	33		0.98 (0.58, 1.66)	9.80
South Africa	1965	27		0.63 (0.39, 1.00)	10.13
Subtotal (I–squared = 20.2%, p = 0.285)				0.81 (0.54, 1.21)	24.81
with estimated predictive interval				. (0.03, 23.28)	
.					
Tropical, < 23.5° latitude					
Puerto Rico	1949	18		0.71 (0.57, 0.89)	11.26
Madanapalle	1950	13		0.80 (0.52, 1.25)	10.25
Haiti	1965	18		0.20 (0.08, 0.50)	7.36
Madras	1968	13		1.01 (0.89, 1.14)	11.50
Subtotal (I–squared = 83.7%, p = 0.000)				0.72 (0.50, 1.04)	40.37
with estimated predictive interval				. (0.15, 3.42)	
.					
Overall (I–squared = 92.1%, p = 0.000)				0.51 (0.34, 0.77)	100.00
with estimated predictive interval				. (0.12, 2.24)	
NOTE: Weights are from random effects analysis					

.1   1   10
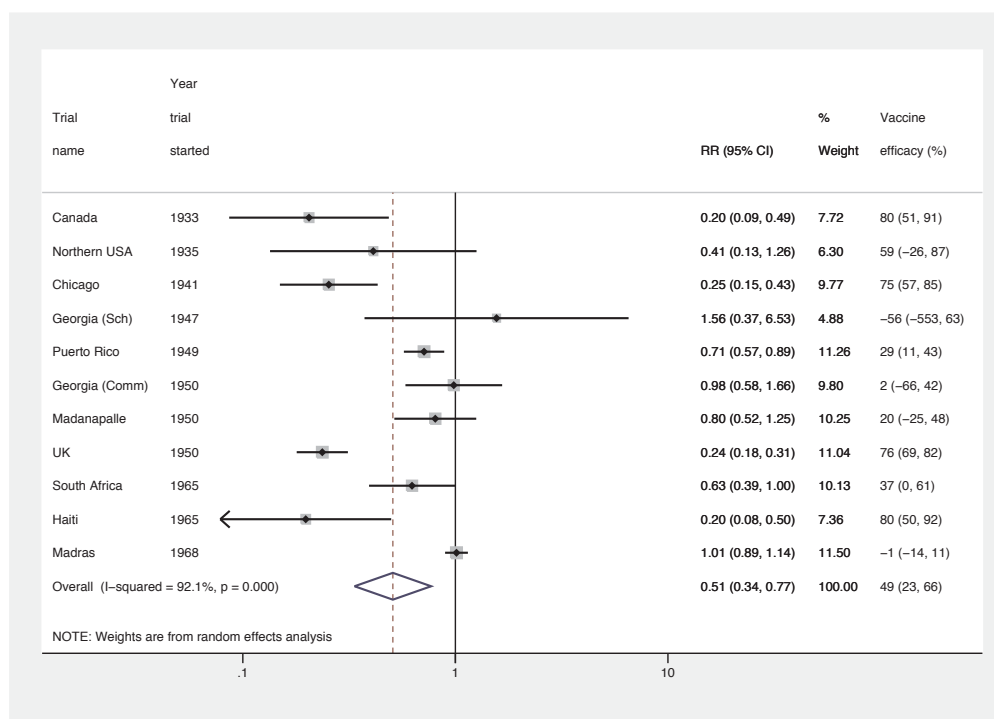
Figure 5. Forest plot displaying a random-effects meta-analysis of the effect of BCG vaccine on incidence of tuberculosis. Results are stratified by latitude region and the prediction interval for a future trial is displayed for each and overall.

◁

9.3 Vaccine efficacy

Results from the analysis of 2×2 data from vaccine trials may be reexpressed as the *vaccine efficacy* (also known as the *relative-risk reduction*); defined as the proportion of cases that would have been prevented in the placebo group had they received the vaccination (Kirkwood and Sterne 2003). The formula is

$$\text{Vaccine efficacy (VE)} = 100\% \times \left(1 - \frac{\text{risk of disease in vaccinated}}{\text{risk of disease in unvaccinated}} \right)$$

$$= 100\% \times (1 - \text{RR})$$

In metan, data are entered in the same way as any other analysis of 2×2 data and the option efficacy added. Results are displayed as odds ratios or RRs in the table and forest plot, but another column is added to the plot showing the results reexpressed as vaccine efficacy.

▷ **Example**

The BCG data are reanalyzed here, with results also displayed in terms of vaccine efficacy. The resulting forest plot is displayed in figure 6.

```
. metan tcases tnoncases ccases cnoncases, rr random efficacy
> lcols(trialnam startyr) textsize(150) notable xlabel(0.1, 10)
```

| Trial name	Year trial started		RR (95% CI)	% Weight	Vaccine efficacy (%)
Canada	1933		0.20 (0.09, 0.49)	7.72	80 (51, 91)
Northern USA	1935		0.41 (0.13, 1.26)	6.30	59 (−26, 87)
Chicago	1941		0.25 (0.15, 0.43)	9.77	75 (57, 85)
Georgia (Sch)	1947		1.56 (0.37, 6.53)	4.88	−56 (−553, 63)
Puerto Rico	1949		0.71 (0.57, 0.89)	11.26	29 (11, 43)
Georgia (Comm)	1950		0.98 (0.58, 1.66)	9.80	2 (−66, 42)
Madanapalle	1950		0.80 (0.52, 1.25)	10.25	20 (−25, 48)
UK	1950		0.24 (0.18, 0.31)	11.04	76 (69, 82)
South Africa	1965		0.63 (0.39, 1.00)	10.13	37 (0, 61)
Haiti	1965		0.20 (0.08, 0.50)	7.36	80 (50, 92)
Madras	1968		1.01 (0.89, 1.14)	11.50	−1 (−14, 11)
Overall (I−squared = 92.1%, p = 0.000)			0.51 (0.34, 0.77)	100.00	49 (23, 66)

NOTE: Weights are from random effects analysis

Figure 6. Forest plot displaying a random-effects meta-analysis of the effect of BCG vaccine on incidence of tuberculosis. Results are also displayed in terms of vaccine efficacy; estimates with a RR of greater than 1 produce a negative vaccine efficacy.

◁

10   More graph options

10.1   metan graph options

Previous users of `metan` may find that they do not like the new box style and prefer a solid black box without the point estimate marker. The option `classic` changes back to this style. There are also options available to change the boxes, diamonds, and other lines. This is achieved by using options that change the standard graph commands that `metan` uses. For instance, the vertical line representing the overall effect may be changed using `olineopt()`, which can take standard Stata *line_options* such as `lwidth()`, `lcolor()`, and `lpattern()`. Boxes are weighted markers and not much can be changed, although shape and color may be modified by using *marker_options* in the `boxopt()` option, such as `msymbol()` and `mcolor()`, or we can dispense with the boxes entirely by using the option `nobox`. The point estimate markers have more flexibility and may also be modified by using *marker_options* in the `pointopt()` option; for instance, labels may by attached to them by using `mlabel()`. The CIs and diamonds may be changed by using *line_options* in the options `ciopt()` and `diamopt()`. For more details, see the `metan` help file and the Stata *Graphics Reference Manual* ([G] **graph**).

▷ **Example**

Here many aspects of the graph are changed and a raw data variable is defined (as in `counts`) and attached to the point estimates in the graph. The resulting graph is not shown here, but a similar application is shown in section 10.3.

```
. gen counts = string(tcases) + "/" + string(tcases+tnoncases) + "," +
> string(ccases) + "/" + string(ccases+cnoncases)

. metan tcases tnoncases ccases cnoncases, rr fixedi second(random) nosecsub
> notable olineopt(lwidth(thick) lcolor(navy) lpattern(dot))
> boxopt(msymbol(triangle) mcolor(dkgreen))
> pointopt(mlabel(counts) mlabsize(tiny) mlabposition(5))
```

◁

10.2   Overall graph options

Any graph options that come under the *overall*, *note*, and *caption* sections of Stata's `graph twoway` command may be added to a `metan` command, and the x axis (and y axis if required) may have a title added. The options `aspect()` or `xsize()` and `ysize()` may be used to specify different aspect ratios (e.g., portrait). The default aspect ratio of a Stata graph is around 0.7 (height/width), and `metan` tries to stick to this shape; although graphs that are more naturally displayed as long or wide will be reshaped to some degree. Use of the above options will control this more precisely.

Finally, the use of schemes is also supported. As colors of boxes and so on are defined within `metan`, these will not always give the desired result but may produce some interesting effects. Try, for example, using the scheme `economist`. More on schemes can be found in [G] **schemes intro**.

10.3   Notes on graph building

It can be useful to declare local or global macros that contain portions of code that are frequently used. For example, if the forest plot always has triangular "boxes" in forest green, contains the same columns of data, and so on, global macros may be declared for these bits of code. These can then be reused for a series of meta-analyses to specify the look and contents of the graphs. These could also be declared in an ado-file so that they are ready to use in every Stata session. This idea is similar to using Stata graph schemes.

▷ **Example**

   Macros are defined to control various aspects of the graph and then used in the `metan` command. The resulting forest plot is displayed in figure 7.

```
. global metamethod rr fixedi second(random) nosecsub
. global metacolumns lcols(trialnam startyr latitude) astext(60)
. global metastyle boxopt(mcolor(forest_green) msymbol(triangle))
> pointopt(msymbol(smtriangle) mcolor(gold) msize(tiny)
> mlabel(counts) mlabsize(tiny) mlabposition(2) mlabcolor(brown))
> diamopt(lcolor(black) lwidth(medthick)) graphregion(fcolor(gs10)) boxsca(80)
. global metaopts favours(decreases TB # increases TB)
> xlabel(0.1, 0.2, 0.5, 2, 5, 10) notable
. metan tcases tnoncases ccases cnoncases,
> $metamethod $metacolumns $metastyle $metaopts by(lat_cat) xsize(10) ysize(8)
```

(Continued on next page)
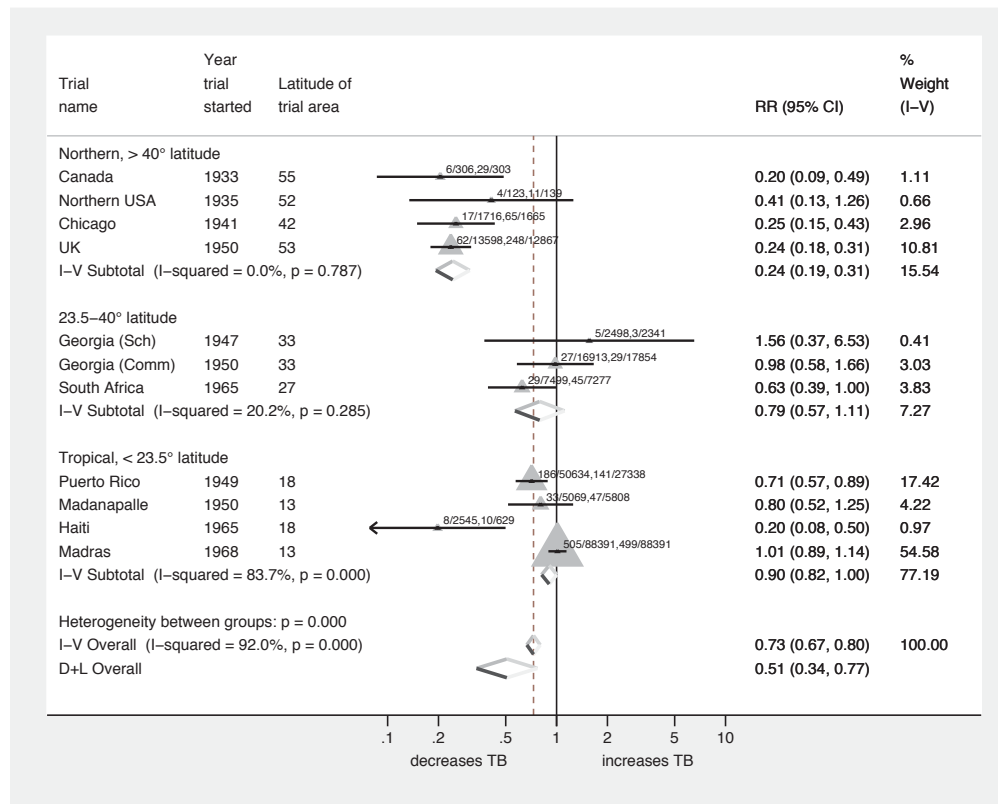
Figure 7. Forest plot displaying an inverse-variance weighted fixed-effect meta-analysis of the effect of BCG vaccine on incidence of tuberculosis. Results are stratified by latitude region, and the overall random-effects estimate is also displayed. Various options have been used to change the display of the graph.

◁

11   Variables and results produced by metan

11.1   Variables generated

When odds ratios (OR) or RRs are combined from 2×2 data and the log option is not used, the SE log-OR or log-RR is saved in a variable named _selogES, to make clear that it is the SE log-OR or RR and not on the same scale. If the log option is used, the standard error is named _seES, as it is on the same scale as the estimate itself. In both cases, the estimate is called _ES.

It is possible to calculate the standard error of ORs and RRs by the delta method; this is what Stata does, for example, with the results reported by the logistic command.

However, the distribution of ratios is in general highly skewed, and for this reason, `metan` does not attempt to record the standard error of either the OR or RR.

Absolute measures (risk differences or mean differences) are symmetric and may be assumed to be normally distributed via the central limit theorem. Here `metan` stores these quantities in `_ES` and their standard errors in `_seES`. The derived variables incorporate the correction for zero cells (see section 8.1).

`_ES`	Effect size (ES)
`_seES`	Standard error of ES
`_selogES`	Standard error of log ES
`_LCI`	Lower confidence limit for ES
`_UCI`	Upper confidence limit for ES
`_WT`	Study percentage weight
`_SS`	Study sample size

11.2   Saved results (macros)

As with many Stata commands, macros are left behind containing the results of the analysis. If two methods are specified by using the option `second()`, some of these are repeated; for example, `r(ES)` and `r(ES_2)` give the pooled-effects estimates for each method. Subgroup statistics when using the `by()` option are not saved; if these are required for storage, it is recommended that a program be written that analyzes subgroups separately (perhaps using the `nograph` and `notable` options).

(Continued on next page)

Name	Second	Description
r(ES)	r(ES_2)	pooled-effect size (if the `log` option is specified with `or` or `rr`, this is the pooled log-OR or log-RR)
r(seES)	r(seES_2)	standard error of pooled-effect size with symmetrical CI, i.e., mean differences, risk difference, log-OR, and log-RR using `log` option
r(selogES)	r(selogES_2)	standard error of log-OR or log-RR when ORs or RRs are combined without the `log` option
r(ci_low)	r(ci_low_2)	lower CI of pooled-effect size
r(ci_upp)	r(ci_upp_2)	upper CI of pooled-effect size
r(z)		Z-value of effect size
r(p_z)		p-value for significance of effect size
r(het)		chi-squared test for heterogeneity
r(df)		degrees of freedom (number of informative studies minus 1)
r(p_het)		p-value for significance of test for heterogeneity
r(i_sq)		the I^2 statistic
r(tau2)		estimated between-study variance (random-effects analyses only)
r(chi2)		chi-squared test for significance of odds ratio (fixed-effect OR only)
r(p_chi2)		p-value for the above test
r(rger)		overall event rate, group 1 (if binary data are combined)
r(cger)		overall event rate, group 2 (see above)
r(measure)		effect measure (e.g., RR, SMD)
r(method_1)	r(method_2)	analysis method (e.g., M-H, D+L)

12   References

Bradburn, M. J., J. J. Deeks, and D. G. Altman. 1998. sbe24: metan – an alternative meta-analysis command. *Stata Technical Bulletin* 44: 4–15. Reprinted in *Stata Technical Bulletin Reprints*, vol. 8, pp. 86–100. College Station, TX: Stata Press.

Breslow, N. E., and N. E. Day. 1980. *Statistical Methods in Cancer Research: Volume I—The Analysis of Case-Control Studies.* Lyon, UK: International Agency for Research on Cancer.

Colditz, G. A., T. F. Brewer, C. S. Berkey, M. E. Wilson, E. Burdick, H. V. Fineberg, and F. Mosteller. 1994. Efficacy of BCG vaccine in the prevention of tuberculosis.

Meta-analysis of the published literature. *Journal of the American Medical Association* 271: 698–702.

Deeks, J. J., D. G. Altman, and M. J. Bradburn. 2001. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In *Systematic Reviews in Health Care: Meta-analysis in context*, ed. M. Egger, G. D. Smith, and D. G. Altman, 285–321.

Fine, P. E. 1995. Variation in protection by BCG: implications of and for heterologous immunity. *Lancet* 346: 1339–1345.

Harbord, R. M., and J. P. T. Higgins. 2008. Meta-regression in Stata. *Stata Journal*. Forthcoming.

Higgins, J. P. T., and S. G. Thompson. 2001. Presenting random effects meta-analyses: where are we going wrong? In *9th International Cochrane Colloquium*. Lyon, France.

———. 2004. Controlling the risk of spurious findings from meta-regression. *Statistics in Medicine* 23: 1663–1682.

Higgins, J. P. T., S. G. Thompson, J. J. Deeks, and D. G. Altman. 2003. Measuring inconsistency in meta-analyses. *British Medical Journal* 327: 557–560.

Kirkwood, B. R., and J. A. C. Sterne. 2003. *Essential Medical Statistics*. 2nd ed. Oxford: Blackwell Science.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS – A Bayesian modelling framework: concepts, structure and extensibility. *Statistics and Computing* 10: 325–337.

Poole, C., and S. Greenland. 1999. Random-effects meta-analyses are not always conservative. *American Journal of Epidemiology* 150: 469–475.

Reis, I., K. F. Hirji, and A. Afifi. 1999. Exact and asymptotic tests for homogeneity in several 2×2 tables. *Statistics in Medicine* 18: 893–906.

Sharp, S., and J. Sterne. 1997. sbe16: Meta-analysis. *Stata Technical Bulletin* 38: 9–14. Reprinted in *Stata Technical Bulletin Reprints*, vol. 7, pp. 100–106. College Station, TX: Stata Press.

Sterne, J. A. C., D. Gavaghan, and M. Egger. 2000. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology* 53: 1119–1129.

Sweeting, M. J., A. J. Sutton, and P. C. Lambert. 2004. What to add to nothing? Use and avoidance of continuity corrections in meta-analysis of rare events. *Statistics in Medicine* 23: 1351–1375.

Warn, D. E., S. G. Thompson, and D. J. Spiegelhalter. 2002. Bayesian random effects meta-analysis of trials with binary outcomes: methods for absolute risk difference and relative risk scales. *Statistics in Medicine* 21: 1601–1623.

About the authors

Ross Harris is a research associate in medical statistics at the Department of Social Medicine, University of Bristol, Bristol, UK. His research interests include meta-analysis, particularly summarizing dose–response relationships from published data and examining sources of bias in randomized trials; epidemiology of HIV and AIDS and methods for dealing with missing data.

Michael Bradburn is a medical statistician in the Health Services Research Unit, University of Sheffield, Sheffield, UK, and wrote most of the original `metan` code. His current research is focused on randomized trials.

Jonathan Deeks is professor of health statistics at the University of Birmingham, Birmingham, UK, and head of the Medical Statistics Group and Diagnostic Evaluation Support Unit in the Department of Public Health. His work has focused on issues related to meta-analysis and more recently diagnostic test evaluations, including both clinical applications and methodological developments. Currently, Jon Deeks is the elected representative for Methods Groups on the Steering Group of the Cochrane Collaboration and is leading the implementation of Reviews of Diagnostic Test Accuracy in the Cochrane Collaboration.

Roger Harbord is a research associate in medical statistics in the Department of Social Medicine, University of Bristol, Bristol, UK. He is a coconvenor of the Cochrane Collaborations Screening and Diagnostic Tests Methods Group.

Douglas Altman is professor of statistics in medicine at the University of Oxford, Oxford, UK, and founding director of the Centre for Statistics in Medicine. His research interests include the use and abuse of statistics in medical research, studies of prognosis, regression modeling, systematic reviews and meta-analysis, randomized trials, reporting guidelines, and studies of medical measurement.

Jonathan Sterne is professor of medical statistics and epidemiology in the Department of Social Medicine, University of Bristol, Bristol, UK. His research interests include statistical methods for epidemiology and health services research, meta-analysis and systematic reviews, clinical epidemiology of HIV and AIDS in the era of antiretroviral therapy, and the epidemiology of asthma and allergic diseases.