# The Stata Journal



The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*, *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

**Subscriptions** are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

**Subscription rates** listed below include both a printed and an electronic copy unless otherwise mentioned.

| U.S. and Canada | | Elsewhere | |
|---|---|---|---|
| 1-year subscription | $ 79 | 1-year subscription | $115 |
| 2-year subscription | $155 | 2-year subscription | $225 |
| 3-year subscription | $225 | 3-year subscription | $329 |
| 3-year subscription (electronic only) | $210 | 3-year subscription (electronic only) | $210 |
| 1-year student subscription | $ 48 | 1-year student subscription | $ 79 |
| 1-year university library subscription | $ 99 | 1-year university library subscription | $135 |
| 2-year university library subscription | $195 | 2-year university library subscription | $265 |
| 3-year university library subscription | $289 | 3-year university library subscription | $395 |
| 1-year institutional subscription | $225 | 1-year institutional subscription | $259 |
| 2-year institutional subscription | $445 | 2-year institutional subscription | $510 |
| 3-year institutional subscription | $650 | 3-year institutional subscription | $750 |

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.



The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, Stata, Stata Press, Mata, Mata, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2012)
**12**, Number 4, pp. 605–622

# Graphical augmentations to the funnel plot to assess the impact of a new study on an existing meta-analysis

Michael J. Crowther
Department of Health Sciences
University of Leicester
Leicester, UK
michael.crowther@le.ac.uk

Dean Langan
Clinical Trials Research Unit (CTRU)
University of Leeds
Leeds, UK
d.p.langan@leeds.ac.uk

Alex J. Sutton
Department of Health Sciences
University of Leicester
Leicester, UK
ajs22@le.ac.uk

**Abstract.** Funnel plots are currently advocated to investigate the presence of publication bias (and other possible sources of bias) in meta-analysis. A previously described augmentation to the funnel plot—to aid its interpretation in assessing publication biases—is the addition of statistical contours indicating regions where studies would have to be for a given level of significance, as implemented in the Stata package `confunnel` by Palmer et al. (2008, *Stata Journal* 8: 242–254).

In this article, we describe the implementation of a new range of overlay augmentations to the funnel plot, many described in detail recently by Langan et al. (2012, *Journal of Clinical Epidemiology* 65: 511–519). The purpose of these overlays is to display the potential impact a new study would have on an existing meta-analysis, providing an indication of the robustness of the meta-analysis to the addition of new evidence. Thus these overlays extend the use of the funnel plot beyond assessments of publication biases. Two main graphical displays are described: 1) statistical significance contours, which define regions of the funnel plot where a new study would have to be located to change the statistical significance of the meta-analysis; and 2) heterogeneity contours, which show how a new study would affect the extent of heterogeneity in a given meta-analysis.

We present the `extfunnel` command, which implements the methods of Langan et al. (2012, *Journal of Clinical Epidemiology* 65: 511–519), and, furthermore, we extend the graphical displays to illustrate the impact a new study has on lower and upper confidence interval values and the confidence interval width of the pooled meta-analytic result. We also describe overlays for the impact of a future study on user-defined limits of clinical equivalence. We implement inverse-variance weighted methods by using both explicit formulas for contour lines and a simulation approach optimized in Mata.

**Keywords:** gr0054, extfunnel, funnel plots, meta-analysis, graphs

 

# 1 Introduction

The funnel plot is now a standard graphical tool for the investigation of publication biases and the extent of heterogeneity in meta-analyses. In its simplest form, a funnel plot is simply an $x$–$y$ scatterplot of the individual study estimates versus some measure of estimate precision and study sample size. Asymmetry in such a plot can be an indication that publication bias is present.

An extensive set of Stata tools has been developed to facilitate the generation of funnel plots; for example, `metafunnel` and `metabias` produce funnel plot displays with various augmentations, such as a line for the pooled effect size. For more details about these commands, see Sterne and Harbord (2004); for further general guidance on the use of graphical tools in meta-analyses, see Anzures-Cabrera and Higgins (2010). An example of a funnel plot of trials of treatment for antidepressant versus placebo is presented in figure 1 (see Moreno et al. [2009] for further details). Here the plot is highly asymmetric, which is indicative of possible publication bias.

```
. use antid
. metafunnel ES seES, noline xtitle("Standardized Effect Size")
> ytitle("Standard Error")
note: default data input format (theta, se_theta) assumed
```

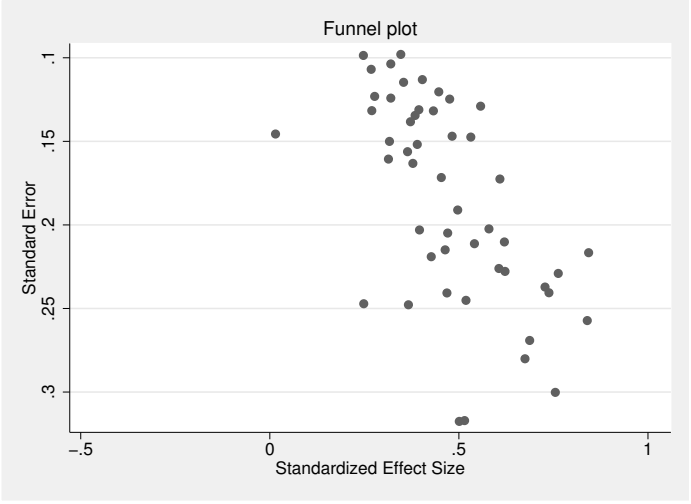


Figure 1. Funnel plot of new-generation antidepressant versus placebo for depression

A recent augmentation to the funnel plot by Peters et al. (2008) is the inclusion of statistical contours indicating regions in which studies would have to be for a given level of statistical significance. This feature is intended to aid the assessment of whether funnel plot asymmetry is likely due to publication biases or other causes. This has also been implemented in Stata as the `confunnel` command by Palmer et al. (2008).

In this article, we present a range of further graphical overlays for the funnel plot, illustrating the potential impact a new study may have when added to an existing meta-analysis. These overlays are similar to the contours for aiding the assessment of publication bias (the previous contours focus on the significance of individual studies, whereas the contours presented here focus on inferences relating to meta-analysis). However, they have a very different purpose that expands the uses of the funnel plot and is broadly applicable across meta-analyses of intervention trials, studies on the accuracy of diagnostic tests, etiological observational studies, etc. The overlays include statistical significance contours, which highlight regions where a new study would have to lie to change the statistical significance of the summary estimate of the present meta-analysis, and heterogeneity contours, which show how a new study would affect the heterogeneity of the meta-analysis. A full description of these overlays, their algebraic derivations (where possible), and applications are available in Langan et al. (2012).

We then extend the displays of Langan et al. (2012) to illustrate how a new study would affect the lower or upper confidence interval and the confidence interval width of the pooled meta-analytic estimate. We also show how these may be of particular interest with regard to diagnostic tests. Furthermore, we adapt the approach to develop a graphical display to assess the impact of a future study on user-defined limits of clinical equivalence described, for example, in Sutton et al. (2007).

The methodology is summarized in section 2, followed by a description of the command syntax in section 3. The use of `extfunnel` is shown in examples in section 4, with some additional features described in section 5. We conclude with a discussion in section 5.

# 2 Methodology

## 2.1 Statistical significance contours

When a study is added to an existing meta-analysis, it is interesting to investigate how this new study affects the statistical significance and direction of the pooled estimate and serves as an indication of the robustness of the original meta-analysis. By choosing plausible ranges for the new study's effect size and associated standard error, we can add each combination to the meta-analysis, which is subsequently re-meta-analyzed, and finally record the statistical significance of the new pooled estimate. By plotting the new study's effect size and standard error, color coded by the now updated meta-analysis' statistical significance, we obtain contour regions illustrating the robustness of the meta-analysis. We describe and illustrate the approach under fixed- and random-effects models using the inverse-variance weighted method.

### Fixed effect

Under a fixed-effects model, explicit formulas for the contours can be derived based on the inverse-variance method. For further details, see Langan et al. (2012).

#### Random effect

Under a random-effects model, each new effect estimate and standard error spanning the entire range of the funnel plot must be combined with the original meta-analysis and analyzed separately to calculate the statistical significance level and direction of the updated pooled estimate. This is because explicit contour lines cannot be derived under a random-effects model. The computational issues of this are discussed in section 3.3.

## 2.2 Heterogeneity contours

A further addition to the standard funnel plot is the overlay of heterogeneity contours. These contours serve to illustrate how the between-study heterogeneity would be affected by the addition of a new study. This can be illustrated using either the between-study variance $\tau^2$ or the $I^2$ statistic.

## 2.3 Alternative targets of inference

Because of the well-documented limitations of focusing exclusively on $p$-values, one can take an approach to inference that focuses more on effect-size estimation. For example, the impact of a new study on the lower or upper limits of the confidence interval of the pooled meta-analytic effect size may be of interest. For instance, in a diagnostic study where accuracy estimation rather than differences between tests is of primary concern, we may be interested in the effect a new study would have on the lower bound of the meta-analytic 95% confidence interval for sensitivity. Similarly, the confidence interval width may be of interest, and how a new study would affect the precision of the pooled estimate can be considered.

## 2.4 Limits of clinical equivalence

An approach to inference that aims to combine clinical information with statistical information is the use of user-specified limits of clinical equivalence. Within these limits, the two interventions are considered equivalent. Although defining the limits is of course subjective, external information can be obtained to inform the defined limits. Eight distinct scenarios are described in figure 2, detailing the characteristics of the confidence interval in relation to the limits of clinical equivalence and providing the subsequent interpretation. Further discussion on the limits of clinical equivalence can be found in Parmar, Ungerleider, and Simon (1996) and Sutton et al. (2007).

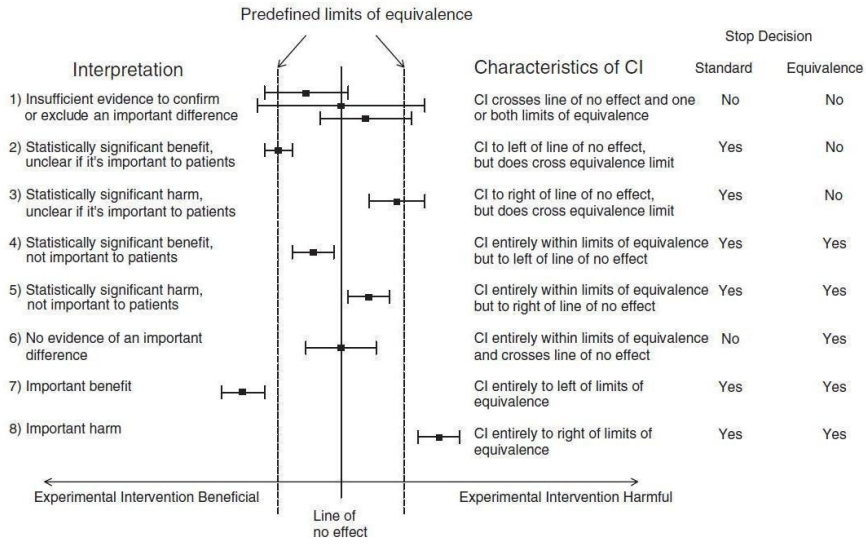


Figure 2. Limits of clinical equivalence (image adapted from Sutton et al. [2007])

# 3 The extfunnel command

## 3.1 Syntax

`extfunnel` *varname1 varname2* [ *if* ] [ *in* ] [ `, fixedi randomi cpoints(#) null(#) isquared(`*numlist*`) tausquared(`*numlist*`) measure(lci|uci|ciwidth) loe(`*numlist*`) loeline newstudycontrol(#) newstudytreatment(#) or rr xrange(`*numlist*`) yrange(`*numlist*`) sumd sumdposition(#) prediction nonullline nopooledline noshading noscatter nometan label([namevar=`*namevar*`], [yearvar=`*yearvar*`]) eform scheme(grayscale|color) addplot(`*string*`) level(#)` *twoway_options* ]

The first variable, *varname1*, must correspond to the effect estimates assumed to be normally distributed (for example, log odds-ratios) with *varname2*, the associated standard errors of the effect estimates. `extfunnel` requires `metan` to be installed.

## 3.2 Options

`fixedi` specifies a fixed-effects model by using the inverse-variance method. This is the default.

`randomi` specifies a random-effects model by using the method of DerSimonian and Laird, with the estimate of heterogeneity being taken from the inverse-variance fixed-effects model.

`cpoints(#)` specifies the number of points to evaluate either the shaded statistical significance contours or the heterogeneity contours. The default numbers for fixed- and random-effects meta-analyses are 3500 and 100, respectively. When a random meta-analysis is invoked, the maximum number of contour points is 500. A larger number of `cpoints()` results in a smoother graph but takes longer to compute (see section 3.3 for more details).

`null(#)` is the value of the null hypothesis for the effect estimate. The default is `null(0)`. This is the value that `lci`, `uci`, or `ciwidth` is compared with when `measure()` is specified. If `lci` or `uci` is specified, the value of `null()` is compared with the lower or upper confidence interval value, respectively, of the updated meta-analyses and is color coded depending on whether the updated estimate is less than or greater than the `null()`. If `ciwidth` is specified, then the width confidence interval of the updated meta-analyses is compared with the value defined by `null()`.

`isquared(`*numlist*`)` specifies the values that define the $I^2$ contours. *numlist* must be of maximum length 5 and should have elements in the range 0–100.

`tausquared(`*numlist*`)` specifies the values that define the between-study variance ($\tau^2$) contours. *numlist* must be a vector of maximum length 5 and should have elements in the range 0–infinity.

`measure(lci | uci | ciwidth)` defines the target of inference, which can be one of `lci`, `uci`, or `ciwidth`.

`loe(`*numlist*`)` defines the limits of clinical equivalence. The default legend assumes a beneficial and detrimental effect in specific directions. The legend can be relabeled using `legend(order(1 "text1" 2 "text2" ...))`. For further details, see Sutton et al. (2007).

`loeline` displays the limits of clinical equivalence.

`newstudycontrol(#)` defines the number of patients in the control arm of a new trial. `newstudycontrol()` and `newstudytreatment()` (explained next) defined together produce a statistical significance contour graph, whereby each possible permutation of results is calculated and analyzed within the appropriate meta-analysis model. Odds ratios and risk ratios are supported.

`newstudytreatment(#)` defines the number of patients in the treatment arm of a new trial.

`or` specifies to use log odds-ratios. This is the default; alternatively, `rr` can be specified for risk ratios. `or` is valid only when `newstudycontrol()` and `newstudytreatment()` are specified.

`rr` specifies to use log risk-ratios. `rr` is valid only when `newstudycontrol()` and `newstudytreatment()` are specified.

`xrange(`*numlist*`)` defines the range of effect estimates used to create the shaded contours.

`yrange(`*numlist*`)` defines the range of standard errors used to create the shaded contours.

`sumd` displays the summary diamond.

`sumdposition(#)` defines the vertical coordinate where the summary diamond should be placed. The default is `sumdposition(0)`.

`prediction` displays a prediction interval for a new trial based on the current meta-analysis. The $y$-axis position is defined by `sumdposition()`.

`nonullline` suppresses the display of the vertical line of no effect.

`nopooledline` suppresses the display of the vertical line at the pooled effect estimate.

`noshading` suppresses the display of shaded regions.

`noscatter` suppresses the display of the scatter of original study effects.

`nometan` suppresses the display of original meta-analysis results using `metan`.

`label(`[`namevar=`*namevar*]`,` [`yearvar=`*yearvar*]`)` labels the variable by its name, year, or both. This is a `metan` option. Either option or both options may be left blank. For the table display, the overall length of the label is restricted to 20 characters.

`eform` exponentiates the $x$-axis labels (valid only when the input variables are log transformed, for example, log odds-ratios or log risk-ratios).

`scheme(grayscale`|`color)` specifies the color scheme of the graph. The default is `scheme(grayscale)`. `scheme(color)` can be useful to distinguish areas when `loe()` is specified.

`addplot(`*string*`)` allows additional twoway plots to be overlayed on the `extfunnel` plot.

`level(#)` specifies the confidence level, as a percentage, for confidence intervals. The default is `level(95)`.

*twoway_options*; see [G-3] ***twoway_options***.

## 3.3 Computational details

As discussed in section 2.1, the inverse-variance weighted method, when applied in a random-effects setting, results in no closed-form expression to calculate the boundary contours between regions of statistical significance for the updated meta-analysis. Using either the default range of the $x$ and $y$ axes or the ranges entered in `xrange()` and `yrange()`, each $x$ and $y$ range is split at $n$ points, where $n$ is defined by `cpoints()`. So for example, `cpoints(500)` would result in $500 \times 500 = 250{,}000$ individual meta-analyses being conducted, with each having its statistical significance calculated. Sets of four adjacent points are then analyzed for the same statistical significance and are color coded into three categories (`Sig. effect < NULL`, `Nonsig. effect`, `Sig. effect > NULL`).

Conducting 250,000 meta-analyses will be computationally intensive. For example, if `metan` was used and specified method `randomi`, with the `nograph` and `notable` options, it would take approximately 24 hours on an Intel Core 2 Duo 3.0 GHz desktop computer to execute the `metan` command and analyze the pooled estimate and confidence interval. For this reason, the inverse-variance weighted method has been implemented in Mata, whereby all $n \times n$ random-effects meta-analyses are conducted simultaneously. When $n = 500$, this reduces computation time to approximately 90 seconds (60 seconds of which involves building the `twoway` graph). Furthermore, attempting to plot thousands of individual `twoway area` graphs on the same graph would cause an overflow of Stata's string limits; therefore the data are prepared row by row into areas of the same statistical significance.

This implementation is also used to create the fixed and random contour graphs when the target of inference is `lci`, `uci`, `ciwidth`, or the limits of clinical equivalence (although closed-form expressions may be obtainable in some contexts). Note that `extfunnel` can be quite memory intensive.

# 4 Example uses of extfunnel

In this section, we detail some features of `extfunnel`.

## 4.1 Statistical significance contours

### Fixed effect

Figure 3 shows a forest plot from a fixed-effects meta-analysis of four trials investigating the change in Epworth score for an oral appliance versus continuous positive airways pressure for treating obstructive sleep apnea. For further details, see Lim et al. (2006).

```
. use epworth
. metan ES seES, fixedi notable texts(220) xlabel(-6, -4, -2, 2, 0, 2, 4, 6)
> xtitle("Effect Size")
```

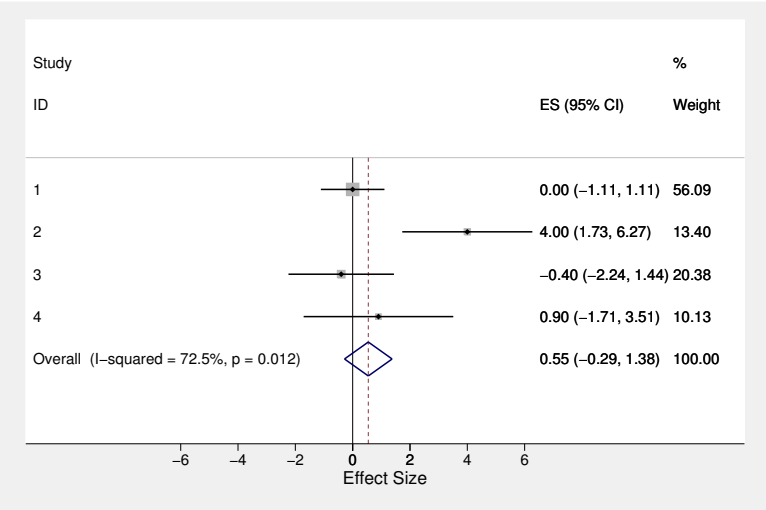


Figure 3. Forest plot of an oral appliance versus continuous positive airways pressure with outcome in the Epworth score

A nonstatistically significant (at the 5% level) summary effect estimate of 0.55 (95% CI; $[-0.29, 1.38]$) was found. By default, `extfunnel` displays the results from a `metan` call on the original meta-analysis. This can be suppressed by specifying the `nometan` option.

```
. use epworth
. extfunnel ES seES
Original meta-analysis results:
           Study     |     ES    [95% Conf. Interval]     % Weight
---------------------+-------------------------------------------------
1                    |  0.000      -1.109     1.109         56.09
2                    |  4.000       1.730     6.270         13.40
3                    | -0.400      -2.240     1.440         20.38
4                    |  0.900      -1.711     3.511         10.13
---------------------+-------------------------------------------------
I-V pooled ES        |  0.546      -0.285     1.376        100.00
---------------------+-------------------------------------------------
  Heterogeneity chi-squared =  10.91 (d.f. = 3) p = 0.012
  I-squared (variation in ES attributable to heterogeneity) =  72.5%
  Test of ES=0 : z=   1.29 p = 0.198
Building graph:
```

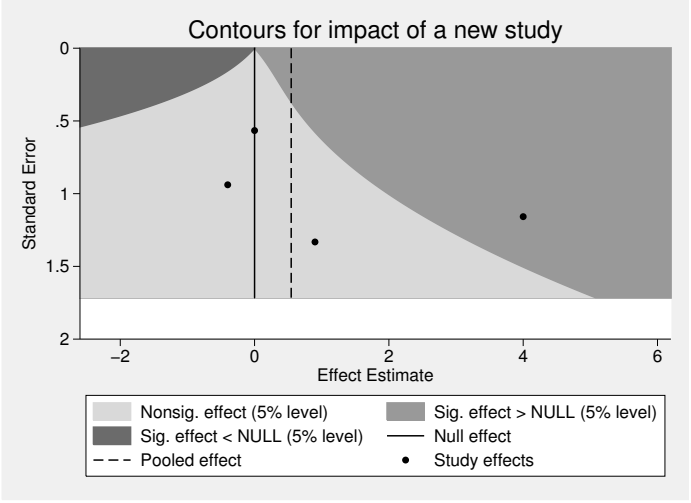


Figure 4. `extfunnel` plot of an oral appliance versus continuous positive airways pressure with outcome in the Epworth score

Using all the default options in this way produces figure 4, which presents contours for areas in which a new study would have to lie for the pooled result to be significant at the 5% level. Regions for significant effects in both directions are present on this plot. It is clear that the existing meta-analysis may not be robust to the impact of a new study. Given that one of the four studies lies in the region where a new study would have to be to change the summary estimate in favor of the continuous positive airways pressure, it is certainly plausible that a new study could alter the conclusion of the existing meta-analysis.

### Random effect

Figure 5 illustrates a meta-analysis of Sanchi versus control for the treatment of ischemic stroke. For further details, see Chen et al. (2008). The log risk-ratios and associated standard errors are combined with a random-effects meta-analysis model. To illustrate the pixel-by-pixel approach, we compare `extfunnel` calls with the default `cpoints(100)` and the maximum `cpoints(500)`. Figure 5 illustrates the improved smoothness of the contours when the maximum `cpoints()` is used. The user-written `grc1leg` is used to combine the two `extfunnel` graphs.

```
. use sanchi
. quietly extfunnel ES seES, randomi xlabel(0.1 0.25 0.5 1 2)
> name(graph1, replace) eform title("100 intervals")
. quietly extfunnel ES seES, randomi cpoints(500) xlabel(0.1 0.25 0.5 1 2)
> eform name(graph2, replace) title("500 intervals")
. grc1leg graph1 graph2
```

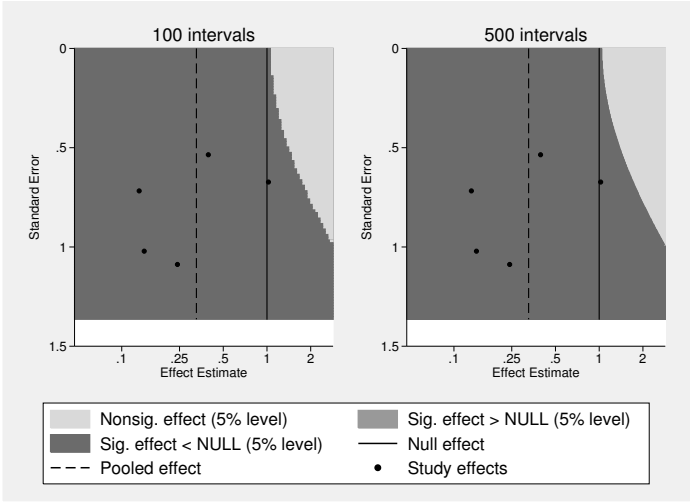


Figure 5. Sanchi versus control for acute ischemic stroke with outcome as proportion of patients with no neurological improvement

Under a random-effects model, the pooled estimate obtained from the original meta-analysis was 0.327 (95% CI; [0.153, 0.699]), with a between-study variance estimate of heterogeneity, $\tau^2 = 0.187$. The original meta-analysis shows a statistically significant reduction in the proportion of patients in the treatment group who had no improvement. From figure 5, it is clear all studies lie in the region of statistical significance with a beneficial treatment effect, which itself dominates the graph area; therefore, the meta-analysis could be considered relatively robust to the addition of a single new trial.

## 4.2 Heterogeneity contours

We detail example plots where we investigate the effect a new study has on estimates of heterogeneity. Heterogeneity statistics include the between-study variance, $\tau^2$, and the $I^2$ statistic. We again use the dataset from Chen et al. (2008) investigating Sanchi versus control for acute ischemic stroke.

### $\tau^2$ contours

Use of the `tausquared()` option can create a contour plot illustrating the effect a new study would have on $\tau^2$. In figure 6, the current estimate of $\tau^2 = 0.187$ is shown, illustrating the current level of heterogeneity in the original meta-analysis. The other contour lines represent combinations of effect estimates and standard errors that would be required in the new study to alter the estimate of $\tau^2$ to each particular value. Up to five distinct contour values of $\tau^2$ can be plotted. In figure 6, we display the summary diamond and a 95% prediction interval (the interval that the underlying effect is estimated to lie within 95% of the time based on the existing meta-analysis) using the `sumd` and `prediction` options, but we suppress the display of the shaded statistical significance contours by invoking the `noshading` option.

```
. use sanchi
. program drop _all
. extfunnel ES seES, tausquared(0.04 0.1 0.187 0.4 0.5) yrange(0 1.5) noshading
> nometan prediction sumd sumdposition(0.2)
Building graph:
```

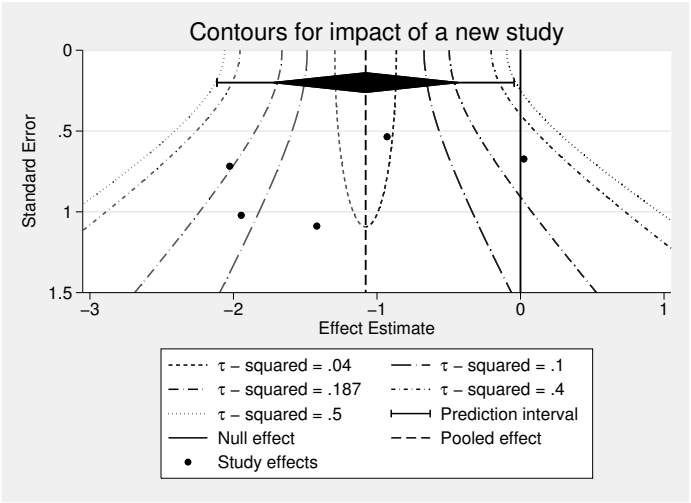


Figure 6. $\tau^2$ heterogeneity contours; Sanchi versus control for acute ischemic stroke

If a new study lies within the region defined by the contours at $\tau^2 = 0.187$, then the estimate of $\tau^2$ would be reduced. Similarly, if a new study lies outside the region, then $\tau^2$ would be increased. If the new study lies on a contour defined by $\tau^2 = 0.5$, then $\tau^2$ would be increased to 0.5 on inclusion of the new study. Mental interpolation between contours can provide a guide to the effect a new study would have on the estimate of $\tau^2$. The prediction interval indicates the likely range of effect sizes (but ignores sampling error) for a new study and thus indicates which region of the plot is most relevant.

Similarly, the effect on the estimates of the $I^2$ statistic can be investigated using `isquared()`. A comparison of the $\tau^2$ and $I^2$ contours, including additional examples, can be found in Langan et al. (2012).

## 4.3 Confidence intervals

Let us now investigate the robustness of the sensitivity of a diagnostic test. The data used in this example come from Geersing et al. (2009), who conducted a meta-analysis examining point-of-care D-dimer tests for detecting venous thromboembolism—more specifically, seven studies evaluating the Clearview Simplify D-dimer test. Analyzing sensitivity (an assessment of specificity could also be conducted using exactly the same approach) under a fixed-effects model produced a pooled sensitivity of 0.853 (95% CI; [0.817, 0.883]). We can investigate the effect of a new study on the lower confidence interval by specifying `measure(lci)` and `null(1.495)`, where invlogit(1.495) = 0.817. Analyses are conducted and plotted on the logit scale.

```
. use simplify
. extfunnel logitsens se_logitsens, fixedi measure(lci) null(1.495) nometan
> xlabel(1 2 3) yrange(0 1) cpoints(500) sumd sumdposition(0.9)
> xtitle("logit(Sensitivity)")
Building graph:
```

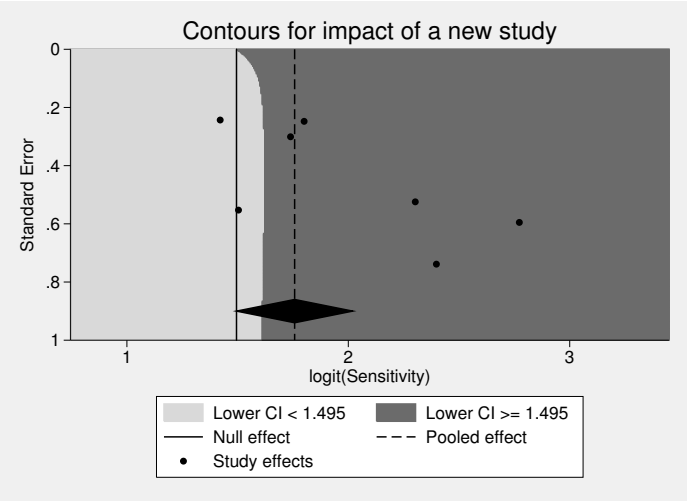


Figure 7. Target of inference—lower confidence interval

From figure 7, we can see where a new study's effect size and standard error would have to lie either to increase or to decrease the lower confidence interval. Similar plots can be produced when the upper confidence interval is of interest or when the confidence interval width is investigated.

## 4.4 Limits of clinical equivalence

We now illustrate how, through the use of the limits of clinical equivalence, `extfunnel` can graphically display a range of scenarios (described in figure 2) for the updated meta-analysis. We illustrate the method using the Epworth score dataset once more, defining the limits of clinical equivalence to be $(-0.25, 0.25)$.

```
. use epworth
. extfunnel ES seES, fixedi loe(-0.25 0.25) nometan sumd sumdposition(1.6)
> yrange(0 2) nopooledline loeline cpoints(400)
Building graph:
```

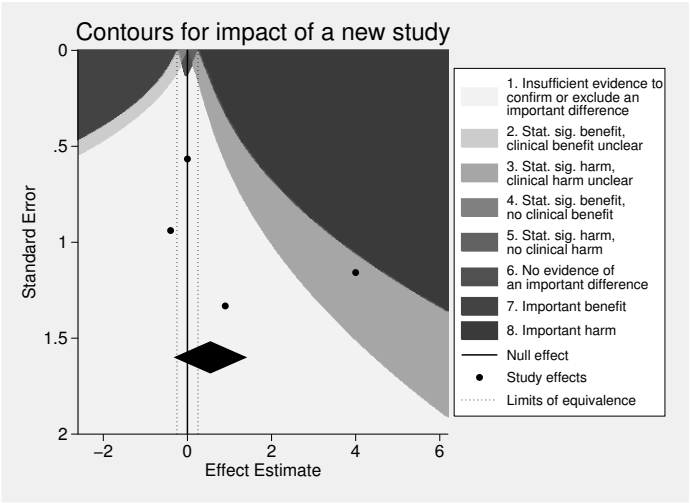


Figure 8. Limits of equivalence

In figure 8, the funnel plot is divided into the eight distinct regions defined in the legend (following the descriptions in figure 2). Here all existing studies lie in either scenario 1 (CI crosses the line of no effect and both limits of equivalence) or scenario 3 (CI lies to the right of the line of no effect but does cross the equivalence line). Given the large proportion of the plot around the pooled estimate that represents scenario 1, changing inferences with one further new study is unlikely unless it is very large (precise) or has an extreme effect size. Note that figure 4 can be considered a special case of figure 8 because figure 8 subdivides further the regions defined in figure 4. We recover figure 4 by grouping areas of the graph where the effect is greater than the null and is statistically significant (areas 3, 5, and 8); by grouping areas where the effect is nonstatistically significant (areas 1 and 6); and by grouping areas where the effect is less than the null and is statistically significant (areas 2, 4, and 7).

# 5 Additional feature

If the measure of interest is an odds ratio or rate ratio, then the effect of a new study with an explicit sample size can be investigated directly through invoking the `newstudytreatment()` and `newstudycontrol()` options. For example, if 100 patients are in both treatment and control arms, we can directly calculate all possible combinations of $2 \times 2$ cell counts, combining each unique new study estimate with the original meta-analysis to produce a graph such as figure 9. This process can be computationally intensive, because it has not been optimized in Mata but relies on `metan`. We use a meta-analysis from a Cochrane review investigating the use of antibiotics versus control to treat the common cold, where the outcome is the alleviation of symptoms within seven days. More details can be found in Arroll and Kenealy (2002). A fixed-effects meta-analysis of the existing studies had a nonstatistically significant pooled odds ratio of 0.796 (95% CI; [0.587, 1.080]).

```
. use colddata
. extfunnel logor selogor, fixedi newstudycontrol(100) newstudytreatment(100)
> nometan eform xlabel(0.01 0.2 1 5 100)
Building graph:
```

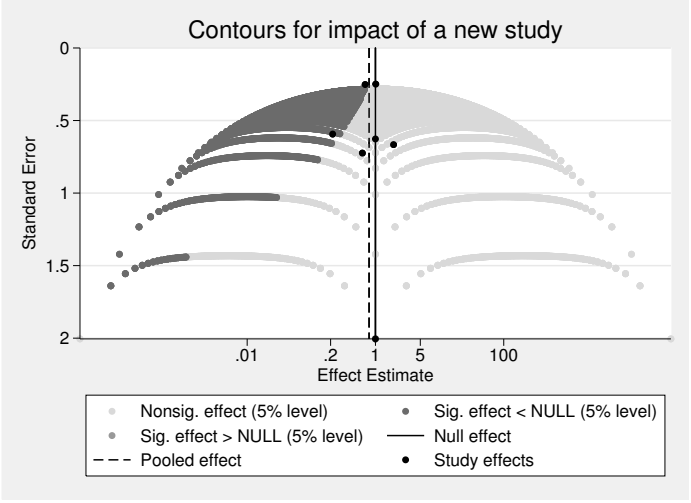


Figure 9. All possible results from the addition of a new study with 100 patients in each arm

Figure 9 shows the possible results, under a fixed-effects model, of combining a trial of size 200 with the original meta-analysis. Given that two of the original studies lie in a region of statistical significance showing a beneficial treatment effect, the original meta-analysis could be considered to be lacking robustness to the influence of a new trial.

# 6 Discussion

We have presented and described the `extfunnel` command, which provides a variety of graphical means of establishing the robustness of a meta-analysis to the inclusion of a new study. The simulation approach implemented in Mata has optimized the re-meta-analyses, providing an efficient and powerful tool encompassing a variety of scenarios.

We envisage that meta-analysts, trialists, and editors of portfolios of systematic reviews will find this display useful when reporting their meta-analyses, designing new studies, and prioritizing updates of existing meta-analyses, respectively (Langan et al. 2012).

The limitations of the approach include no accounting for change in the statistical model. We assume that the original meta-analysis is analyzed with the same fixed or random framework as the updated meta-analysis. Furthermore, only inverse-variance weighted methods are available because, for example, the Mantel–Haenszel approach is prohibitive: it takes into account all cell frequencies in a study's $2 \times 2$ contingency table, which cannot be displayed on a two-dimensional graph. Finally, the plots only consider the inclusion of one further study; an approach that could display the impact of multiple studies would be welcomed.

# 7 Acknowledgments

We thank Julian Higgins for helpful feedback on the program and the article, and Rob Herbert and Manuela Ferreira for invaluable suggestions for improvement and testing of the command. Michael Crowther is funded by a National Institute for Health Research methodology fellowship (RP-PG-0407-10314).

# 8 References

Anzures-Cabrera, J., and J. P. T. Higgins. 2010. Graphical displays for meta-analysis: An overview with suggestions for practice. *Research Synthesis Methods* 1: 66–80.

Arroll, B., and T. Kenealy. 2002. Antibiotics for the common cold. *Cochrane Database of Systematic Reviews* 3: CD000247.

Chen, X., M. Zhou, Q. Li, J. Yang, Y. Zhang, D. Zhang, S. Kong, D. Zhou, and L. He. 2008. Sanchi for acute ischaemic stroke. *Cochrane Database of Systematic Reviews* 4: CD006305.

Geersing, G. J., K. J. M. Janssen, R. Oudega, L. Bax, A. W. Hoes, J. B. Reitsma, and K. G. M. Moons. 2009. Excluding venous thromboembolism using point of care D-dimer tests in outpatients: A diagnostic meta-analysis. *British Medical Journal* 339: b2990.

Langan, D., J. P. T. Higgins, W. Gregory, and A. J. Sutton. 2012. Graphical augmentations to the funnel plot assess the impact of additional evidence on a meta-analysis. *Journal of Clinical Epidemiology* 65: 511–519.

Lim, J., T. J. Lasserson, J. Fleetham, and J. Wright. 2006. Oral appliances for obstructive sleep apnoea. *Cochrane Database of Systematic Reviews* 1: CD004435.

Moreno, S. G., A. J. Sutton, E. H. Turner, K. R. Abrams, N. J. Cooper, T. M. Palmer, and A. E. Ades. 2009. Novel methods to deal with publication biases: Secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *British Medical Journal* 339: b2981.

Palmer, T. M., J. L. Peters, A. J. Sutton, and S. G. Moreno. 2008. Contour-enhanced funnel plots for meta-analysis. *Stata Journal* 8: 242–254.

Parmar, M. K. B., R. S. Ungerleider, and R. Simon. 1996. Assessing whether to perform a confirmatory randomized clinical trial. *Journal of the National Cancer Institute* 88: 1645–1651.

Peters, J. L., A. J. Sutton, D. R. Jones, K. R. Abrams, and L. Rushton. 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology* 61: 991–996.

Sterne, J. A. C., and R. M. Harbord. 2004. Funnel plots in meta-analysis. *Stata Journal* 4: 127–141.

Sutton, A. J., N. J. Cooper, D. R. Jones, P. C. Lambert, J. R. Thompson, and K. R. Abrams. 2007. Evidence-based sample size calculations based upon updated meta-analysis. *Statistics in Medicine* 26: 2479–2500.

### About the authors

Michael Crowther is a research assistant and part-time PhD student in medical statistics who maintains an active research interest in meta-analysis.

Dean Langan is a medical statistician in clinical trials with an active interest in meta-analysis.

Alex Sutton is a professor of medical statistics with a longstanding research interest in evidence synthesis.