BMJ

BMJ 2013;346:f1342 doi: 10.1136/bmj.f1342 (Published 1 March 2013)

ENDGAMES

STATISTICAL QUESTION

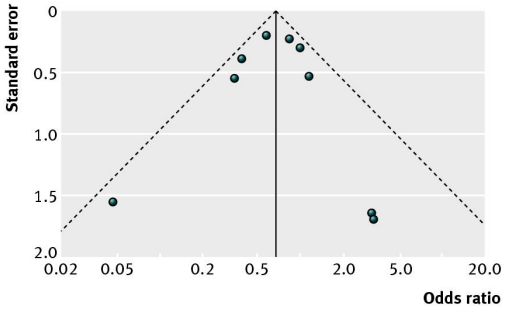
Meta-analyses: how to read a funnel plot

Philip Sedgwick *reader in medical statistics and medical education*

Centre for Medical and Healthcare Education, St George's, University of London, Tooting, London, UK

Researchers undertook a meta-analysis to evaluate the relative short term safety of carotid endarterectomy compared with carotid artery stenting.[1] Randomised controlled trials were included if they compared carotid endarterectomy with carotid artery stenting in patients with carotid artery stenosis (with or without symptoms). In total, 10 trials were identified that reported short term outcomes. The primary endpoint was mortality or stroke within 30 days of the procedure.

The total overall periprocedural risk of mortality or stroke was lower for carotid endarterectomy than for carotid artery stenting (odds ratio 0.67, 95% confidence interval 0.47 to 0.95; P=0.025). A funnel plot was presented (figure). Egger's test gave a P value equal to 0.932.



Funnel plot for the meta-analysis of the short term safety (periprocedural mortality or stroke) of carotid endarterectomy compared with carotid artery stenting

Which of the following statements, if any, are true?

a) The funnel plot is used to detect bias in trials included in the meta-analysis

b) Publication bias will result in asymmetry of the funnel plot

c) The result of Egger's test indicates that bias existed in the trials included in the meta-analysis

Answers

Statements *a* and *b* are true, whereas *c* is false.

The meta-analysis investigated the relative short term safety of carotid endarterectomy compared with carotid artery stenting. Ten trials were identified in total, and for each the odds ratio of periprocedural mortality or stroke for carotid endarterectomy compared with carotid artery stenting was obtained. Each sample odds ratio is an estimate of the population parameter. However, the meta-analysis might not have identified all the relevant trials that had been conducted, which might mean that the total overall estimate was biased. The purpose of the funnel plot and Egger's test was to detect possible bias in the trials that were identified and included in the meta-analysis (*a* is true)

Failure to include in a meta-analysis all of the relevant studies that have been conducted is often, wrongly, attributed solely to publication bias. Other possible sources of bias exist, and these biases are described collectively as reporting bias. Publication bias—the failure to include all relevant trials because they were not published and were therefore not accessible—is the most well known of these biases. Publication of studies depends on the nature and direction of their results and is more likely if the results are significant or perceived as important. Sometimes publication is influenced by the study size, funding source, or research group.

Further examples of reporting bias are language and citation biases. Language bias is the selective inclusion of studies that are published in an easily accessible language, typically English. When identified, the reference lists of these studies are examined for other possible trials. This may result in citation bias—the tendency for those studies most often cited to be identified and included in the meta-analysis.

The funnel plot is a scatter plot of each of the estimated effects for the trials identified, with the sample odds ratio of periprocedural risk of mortality or stroke comparing treatments on the horizontal axis, against the standard error of the estimated effect on the vertical axis. The standard error provides a measure of the precision of the odds ratio as an estimate of the population parameter. Typically, trials with smaller sample sizes produce less precise estimated effects. As sample size increases the precision of the estimated effect increases and the size of the standard error decreases. The vertical axis in the funnel plot is inverted, with zero at the top. Therefore, studies with less precise

p.sedgwick@sgul.ac.uk

estimated effects scatter more widely at the bottom of the plot. It is expected that the estimated effects will scatter uniformly around the total overall estimate of the meta-analysis (represented by the vertical line in the figure) because of sampling error in the selection of samples from the population. As sample size increases, the precision of the estimated effects increases and the spread of points narrows. Therefore, the scatter plot should resemble a funnel.

Sometimes the reciprocal of the standard error is plotted on the vertical axis, and then the axis does not need to be inverted for the points to resemble a funnel. Alternative measures of precision of the estimated effects are sometimes used instead of standard error, including the reciprocal of the sample size or variance of the estimated effect. On the funnel plot in the example above, lines were superimposed to resemble the limits of the predicted funnel shape in the estimated effects of the odds ratio.

In the absence of reporting bias a funnel plot will be symmetrical in shape—that is, the points will be scattered in the shape of a funnel centrally around the total overall estimated effect. If publication bias is present, as for any reporting bias, the plot will be asymmetrical (*b* is true). Although the plot may detect bias, it is not possible to identify which biases are present. Assessment of symmetry in the funnel plot is often subjective. Inspection of the funnel plot in the example above does not suggest asymmetry because the estimated effects are scattered within the superimposed limits. However, assessment is difficult because the number of trials is not large. In general, funnel plots are thought to be unreliable methods of investigating publication bias, particularly if the number of studies is small (less than 10).

Formal statistical tests exist for detecting asymmetry in a funnel plot, including Egger's test. The null hypothesis for Egger's test is that symmetry exists in the funnel plot, with the alternative indicating that asymmetry is present. The P value for Egger's test for the example above is 0.932, so there was no evidence to reject the null hypothesis in favour of the alternative at the 5% level of significance, and it can be concluded that symmetry exists in the funnel plot (*c* is false). Therefore no apparent bias exists in the studies included in the meta-analysis.

Asymmetry in the funnel plot may also have occurred because of poor methodological design in the trials—for example, failure to conceal the allocation process. It was not possible to conceal treatment allocation from patients in any of the trials in the above meta-analysis and this may have led to response bias. Poor methodological design is typically a problem in trials with small sample sizes because it can lead to spuriously inflated estimated treatment effects. This will lead to an absence of studies on the left hand side at the base of the funnel, resulting in asymmetry in the funnel plot.

Competing interests: None declared.

1 Meier P, Knapp G, Tamhane U, Chaturvedi S, Gurm HS. Short term and intermediate term comparison of endarterectomy versus stenting for carotid artery stenosis: systematic review and meta-analysis of randomised controlled clinical trials. *BMJ* 2010;340:c467.

Cite this as: *BMJ* 2013;346:f1342