Considerations and recommendations for figures in Cochrane reviews: graphs of statistical data

4 December 2003 (updated 27 February 2008)

1 Introduction

RevMan can perform and display meta-analyses of dichotomous data, continuous data and 'O – E' statistics and estimates and standard errors (Deeks 2001, Higgins 2008). It also allows graphical displays to be imported that have been created using different software. The purpose of this document is to provide recommendations from the Statistical Methods Group (SMG) of the Cochrane Collaboration regarding the content of graphical displays. It is intended to cover forest plots as displayed by RevMan and additional figures that review authors may wish to include in a Cochrane review.

1.1 Graphs and Cochrane reviews

The purpose of a graph is to present numerical data in visual form. Graphs enable the identification of overall patterns, correlations and outlying observations that might be overlooked in tables of data. Graphs are especially valuable when a table is not an option (for example, presenting numerous data in a scatter diagram) and/or where there is some possible trend to look for. They can save the reader considerable time and effort in absorbing the findings of a systematic review, and can facilitate the comparison of data across different scenarios. However, if poorly designed they can frustrate and even mislead the reader.

There are many ways of analysing and displaying data arising from a systematic review, a meta-analysis or indeed a single study included in a systematic review. Graphical displays for meta-analysis have been discussed by Galbraith (Galbraith 1988), Light *et al* (Light 1994), Pettiti (Petitti 1994) and Sutton *et al* (Sutton 1998). It is expected that the majority of figures deemed appropriate for inclusion in Cochrane reviews will be forest plots. Facilities for drawing forest plots are available within Cochrane review-writing software, and these should be used in preference to other facilities whenever possible.

This document has been developed by members of the SMG to address the following:

- General considerations and recommendations for graphs in systematic reviews
- Recommendations and examples for forest plots
- Recommendations and examples for the following types of plots that might, on occasion, be appropriately included in Cochrane reviews as additional figures
  - Summary forest plots
  - Funnel plots

  - Relationship between intervention effect and a single covariate (meta-regression)
  - Graphical displays particular to dichotomous outcome data (L'Abbé plots and plots relating intervention effect to 'underlying risk')
- Considerations for the following plots that are not specifically encouraged in Cochrane reviews
  - Galbraith (radial) plots
  - Relationship between intervention effect and two or more covariates (meta-regression)
  - Survival curves
  - Cumulative meta-analysis
  - Other graphical displays

The SMG has developed recommendations as guidelines and not as rules. On occasion there may be good reason to approach a graph differently. Further, the types of graph addressed in this document are not a comprehensive list of those that may usefully be included in a systematic review. Given the almost limitless possibilities available to a review author, we place high emphasis on the following general recommendation.

General recommendations

1.1. Every graphical display of data should be assessed by a statistician as part of the editorial process within the relevant Cochrane Review Group, before being submitted as part of a Cochrane review. The assessment should cover appropriateness, clarity and obvious errors. Ideally it should also cover correctness of the data and/or analyses being presented. Establishing correctness of data may require examination of original reports from the included studies.

1.2. Data represented in a graph should be tabulated whenever it is reasonable to do so (this may not be suitable for scatter plots, for example).
Such data may appear within the graph, or elsewhere such as in 'Other data' tables or 'Additional tables' within the Cochrane review.

2 Principles of graphing data

Five principles, discussed in detail by Cleveland (Cleveland 1994), provide a useful framework for creating, selecting or refining a graph. They are (*i*) accuracy, (*ii*) simplicity, (*iii*) clarity, (*iv*) appearance, and (*v*) a well-defined structure. A review author or statistician creating graphs for inclusion in a Cochrane review should also remember that a high proportion of the readership have had no training in research methods or statistics.

There are certain criteria that all graphical displays of data should fulfil. The list below represents an ideal, and incorporates advice drawn from various external sources (Arkin 1940, Simmonds 1980, Schmid 1983, Cleveland 1994). It may not be possible for a review author to control all of these aspects within their chosen software.

Recommendations for all graphical displays

Titles, captions and scales

2.1. The graph should be supplied with a brief, comprehensive title. It may be helpful to supplement this with a caption, that is a sentence or two to aid understanding and interpretation of the picture. The graph, along with its associated title and caption should generally be understandable outside the context of the rest of the document.

2.2. Explanatory variables (variables used to 'predict' changes in other variables) should be on the horizontal axis. This general rule is not followed in some common representations of meta-analysis, and we discuss it further in the context of specific graph types below.

2.3. Every axis should be labelled, identifying both the quantity and its units (using SI units where applicable).

2.4. Ranges of scales should be chosen so that all (or nearly all) the range of the data is included, and so as to maximise use of available space. However, they should not be chosen so that unimportant variation is exaggerated.

2.5. Excluded data (through curtailing axes or other reasons) should be mentioned in a caption to the graph.

2.6. It is generally desirable but not always necessary that key reference values are included on an axis (for example, 0 for a difference measure of intervention effect; 1 for a ratio measure of intervention effect, 0% and 100% for percentages)

2.7. If two or more graphs are to be compared directly (e.g. for subgroups), identical scales should be used.

2.8. There should not be an excessive number of tick marks or gridlines, and these should not interfere with data.

2.9. Sufficient tick marks should be labelled to allow the reader to interpolate values between them. There should be at least 3 tick marks on any axis. A '0.' should be placed in front of decimal points.

2.10. When a log scale is used, the tick marks should be labelled on the original (un-logged scale)

2.11. A reference line should be considered for an important value (for example, a meta-analysis result), though such a line should not interfere with other components of the graph.

Representing data

2.12. The data should stand out so that main trends can be seen at a glance. Superfluous contents should be removed.

2.13. The weight (or thickness) of lines for data should be equal to, or exceed, that for the axes.

2.14. Clear and prominent symbols should be used to show data. Different plotting symbols should be distinguishable, especially if they may overlap.

2.15. Notes or keys should be used to define the meaning of different styles of lines or symbols. Direct labelling of lines or symbols is preferable. Notes and keys may be placed inside or outside the graphing area or within the caption. They should be placed inside the graphing area only when they do not interfere with data or clutter the graph.

2.16. It is important that variability and uncertainty are fully expressed when presenting results, but care must be taken when providing this information on a graph. Error bars may cause confusion

or obscure the main data. Some possibilities are to present variability or uncertainty in separate tables; to use different sized plotting symbols; to extend error bars to one side only; or to plot points off-centre so that error bars do not overlap. All representations of variability or uncertainty must be explained, stating exactly which quantity (for example, standard error, weight, X% confidence interval) is being illustrated.

Perseverance of information

2.17. Graphs (including text within them) should be robust to reproduction and reduction. In particular, information must not be lost if the graph is reproduced in black and white. Whereas colour may be used to enhance the appearance of a graph, it must not be relied upon to distinguish different components.

2.18. Use of different line types can enhance visual impact.

2.1 Principles of meta-analysis

Two of the principles underlying meta-analysis of healthcare intervention studies are as follows.

1. **Compare like with like**. Since studies are undertaken in different populations often using different variations of interventions, with different definitions of outcomes and using different designs, it is appropriate for experimental and control groups to be compared *within* studies and not *across* studies. The within-study comparisons ('intervention effects', or 'effect sizes') are combined across studies in the meta-analysis.

2. **Not all studies are of equal importance.** The amount of weight awarded to each study in a meta-analysis reflects the amount of information in the study.

In using graphical methods for presenting meta-analyses, one would therefore generally expect that

(i) *studies* (rather than, say, patients, interventions or single arms of studies) will be the unit of interest (the points being plotted); and

(ii) the amount of information contained in each study will be reflected in the graph.

When creating graphical displays that are not addressed in this document, it may be helpful to bear these considerations in mind.

3 Forest plots

Forest plots are also known as confidence interval plots. More informal terms include 'blocks and lines plots' and 'blobbograms'. They are the standard means of presenting results of individual studies and meta-analyses (Egger 1997a, Lewis 2001). A forest plot displays results (that is, estimates of intervention effect) and confidence intervals for individual studies and/or meta-analyses. Graphs produced by RevMan are forest plots. An example is given in Figure 1. Each study is represented by a square at the point estimate of intervention effect and a horizontal line extending either side of the block. The area of the block is proportional to the weight assigned to that study in the meta-analysis, and the horizontal line gives a confidence interval (with specified level of confidence). The area of the block and the confidence interval convey similar information, but both have important contributions to the graph. The confidence interval provides a range of intervention effects compatible with the study's result. If it does not pass through the line of no effect this indicates that the result was

individually statistically significant. The size of the block draws the eye towards the studies with larger weight (smaller confidence intervals). Failure to use this second device may result in unnecessary attention to those smaller studies with wider confidence intervals that put more ink on the page (or more pixels on the screen).

Forest plots may include meta-analyses, normally at the bottom of the graph. A variety of methods is available for conducting the meta-analysis, including both classical and Bayesian methods. Forest plots for Bayesian (or empirical Bayes) meta-analyses may include both the original and 'shrunk' estimates of intervention effect for each study. These would normally appear together.

It is conventional to represent all information relevant to each study (or meta-analysis) within a row. This means the horizontal axis of the graph denotes the size of intervention effect (the outcome, or dependent variable). This convention breaks the general rule that independent variables be plotted along the horizontal axis, and several authors (mainly statisticians) have thus drawn such graphs the other way round (Bailey 1987). However, we believe that the break with the general rule is justified, and offers advantages, for the following three reasons. We therefore incorporate the convention into our recommendations.

The 'study' axis is not a numerical scale, so the recommendation is of lesser importance. There is also a 'natural break' between a list of studies and a meta-analytic summary, which may be visually clearer when they are plotted one above the other.

The convention enables written details of each study to be presented alongside the results. As a minimum, an identifier for the study (such as its Study ID) can be included without resorting to vertical or inclined text. Other information such as raw data, study characteristics and the numerical results being plotted may also be presented.

The convention complements the typical presentation of tables of studies, in which studies appear in rows, and characteristics (or results) in columns.

Recommendations for forest plots

3.1. If a forest plot may appropriately be drawn using RevMan, it should be. All remaining recommendations are consistent with forest plots drawn using RevMan.

3.2. Forest plots should be referred to as 'forest plots' in preference to other names.

3.3. The effect measure should be along the horizontal axis.

3.4. Ratio measures of intervention effect (such as odds ratios, relative risks, hazard ratios and rate ratios) should be plotted on the log scale. The labels on the axis, however, should be on the original (anti-logged) scale (Galbraith 1988).

3.5. A reference line should be drawn at the position of no intervention effect.

3.6. Another, usually dashed, line can be added to indicate the estimated pooled effect

3.7. Intervention effect estimates and confidence intervals should be plotted for each study and each meta-analysis.

3.8. The level of confidence for confidence intervals should be stated (for example, 95%, 99%). The levels of confidence need not be the same for individual studies and overall effect, though any differences must be clearly labelled.

3.9. The directions of effect should be clearly shown, preferably directly below the plot (for example, 'Favours aspirin ←' and '→ Favours placebo' *or* 'Aspirin better ←' and '→ Aspirin worse').

3.10. Intervention effect estimates and confidence intervals, or results sufficient to calculate these, must be presented numerically somewhere in the review.

Individual studies

3.11. The size of the block representing a point estimate from a study should usually relate to the amount of information in the study. If a meta-analysis is included, that information should be the weight apportioned to the study in the meta-analysis. If no meta-analysis is included, that information may be the weight that *would* be apportioned to that study in a meta-analysis, or the total sample size in the study. Note that weights depend not only on sample size, but also on the choice of effect measure. (Thus, for example, relative weights are different on the odds ratios scale compared with the risk difference scale).

3.12. It should be possible to identify from which trial each result belongs. This will normally be achieved by including the 'Study ID' alongside the result.

3.13. Additional information such as the summary data and/or the numerical results being plotted can be helpful (Light 1994). This information is presented by default on meta-analyses generated using RevMan (see Figure 1).

3.14. The minimum number of studies appropriate for display in a forest plot is 2. In rare cases the number of studies will be very large, so that the plot cannot be read properly. It may be helpful to present a summary forest plot (see below).

3.15. Studies should have a meaningful order. Often this is alphabetical by study identifier, or according to date of publication. However, it may be helpful to order by some other characteristic, such as duration or dose of treatment.

Meta-analyses

3.16. The method used to perform a meta-analysis should be stated in the plot, in the title or in the caption. For example, it should be clear whether a fixed effect or random effects model has been used.

3.17. If both meta-analyses and individual studies are plotted, a meta-analysis should be plotted in a different style. For example, using a diamond (stretching the width of the confidence interval), or using an unfilled block (with accompanying confidence interval line).

3.18. If a meta-analysis is considered to be inappropriate, unhelpful, misleading or erroneous it should not be included in a forest plot.

4 Summary forest plots

Forest plots may also be used to illustrate results of meta-analyses in the absence of individual study results, for example to enable the comparison of different outcomes, subgroup analyses or sensitivity

analyses (see Figure 2). This is a particularly useful form of graph, and we propose the name 'summary forest plot' to indicate that the individual points represent meta-analyses rather than studies.

Recommendations for summary forest plots

4.1. Recommendations 3.1 to 3.10 for forest plots, and 3.16 to 3.18 for meta-analyses within forest plots, should be followed.

4.2. The review author should consider carefully whether points should be drawn with equally sized blocks, or blocks according to total weight in each meta-analysis. For subgroup analyses and sensitivity analyses, block sizes according to total weight are recommended. When meta-analyses of different outcomes are presented in the same plot it may be more appropriate to use equally sized blocks.

5 Funnel plots

Funnel plots, introduced by Light and Pillemer (Light 1994) and discussed in detail by Egger and colleagues (Egger 1997b, Sterne 2001a), are useful adjuncts to meta-analyses. A funnel plot is a scatter plot of intervention effect against a measure of study size. It is used primarily as a visual aid to detecting bias or systematic heterogeneity. A symmetric inverted funnel shape arises from a 'well-behaved' data set, in which publication bias is unlikely. An asymmetric funnel indicates a relationship between intervention effect and study size. This suggests the possibility of either publication bias or a systematic difference between smaller and larger studies ('small study effects'). Asymmetry can also arise from use of an inappropriate effect measure. Whatever the cause, an asymmetric funnel plot leads to doubts over the appropriateness of a simple meta-analysis and suggests that there needs to be investigation of possible causes.

A variety of choices of measures of 'study size' is available, including total sample size, standard error of the intervention effect, and inverse variance of the intervention effect (weight). Sterne and Egger have compared these with others, and conclude that the standard error is to be recommended (Sterne 2001b). When the standard error is used, straight lines may be drawn to define a region within which 95% of points might lie in the absence of both heterogeneity and publication bias (Sterne 2001b).

In common with confidence interval plots, funnel plots are conventionally drawn with the effect measure on the horizontal axis, so that study size appears on the vertical axis, breaking with the general rule. Since funnel plots are principally visual aids for detecting asymmetry along the intervention effect axis, this makes them considerably easier to interpret. We therefore feel this is justifiable and to be recommended. An example of a funnel plot appears in Figure 3. Funnel plots can be drawn within RevMan.

Recommendations for funnel plots

5.1. The intervention effect measure should be along the horizontal axis.

5.2. Ratio measures of intervention effect (such as odds ratios, relative risks, hazard ratios and rate ratios) should be plotted on the log scale. The ticks and labelled values on the axis, however, should be on the original (anti-logged) scale.

5.3. The measure of study size (on the vertical axis) should generally be the standard error of the intervention effect estimate. A trick to invert the graph so that bigger trials appear at the top is to plot the negative standard error and override (or edit) the axis labels to remove the minus signs (Sterne 2001b).

5.4. Points should all be the same size, since the size of a study is already described using the vertical axis.

5.5. 95% limit lines may be included. If so they should usually be centred around a fixed effect meta-analysis.

5.6. Funnel plots may not be useful for small numbers of studies (for example, a small study effect may difficult to spot among fewer than ten studies)

5.7. Intervention effect estimates and their standard errors, or results sufficient to calculate these, must be presented numerically somewhere in the review.

6 Relationship between intervention effect and a single covariate (meta-regression)

It has been argued that sources of heterogeneity in a meta-analysis should be investigated (Thompson 1994). Often a source of heterogeneity can be summarized as a trial-level covariate, that is some varying characteristic of the trials. A scatter plot with the covariate along the horizontal axis and the intervention effect along the vertical axis provides a convenient visual impression of the relationship (Thompson 2001). Such scatter plots have commonly followed the convention of plotting the covariate (explanatory variable) along the horizontal axis and the intervention effect (outcome variable) on the vertical axis.

Meta-regression is the statistical analysis of the association between intervention effect and the value of one, or more, trial-level covariate(s). The analysis yields a regression line that may be superimposed on the scatter plot. A particular application is when the intervention affects a continuous surrogate endpoint, such as blood pressure or serum cholesterol, in which case it may be hypothesized that the benefit of intervention, say on mortality, would be related to the success in modifying the surrogate. An example of a meta-regression analysis appears in Figure 4.

Recommendations for single variable 'meta-regression' plots

6.1. The covariate (trial-level characteristic) should be along the horizontal axis.

6.2. The intervention effect should be up the vertical axis.

6.3. A reference line at the position of no intervention effect may be useful.

6.4. Ratio measures of intervention effect (such as odds ratios, relative risks, hazard ratios and rate ratios) should be plotted on the log scale. The labels on the axis, however, should be on the original (anti-logged) scale.

6.5. Points should be of a size proportional to weight or trial size (preferably weight).

6.6. Trial weights or sample sizes should not be illustrated using confidence intervals alone (these draw attention to trials with small weights rather than those with large weights).

6.7. A meta-regression line may be plotted.

6.8. Confidence or prediction lines either side of the meta-regression line may be useful. Note that these are unlikely to be parallel to the meta-regression line.

6.9. For dichotomous outcome data, plots of intervention effect against underlying risk (as measured by observed control group event rate) is usually misleading and should be avoided (see below).

6.10. Intervention effect estimates, their standard errors and the covariate values, or results sufficient to calculate these, must be presented numerically somewhere in the review.

7 Graphical displays particular to dichotomous outcome data

7.1 L'Abbé plots

Results of multiple clinical trials with dichotomous outcomes may be represented in a L'Abbé plot, after a paper by L'Abbé and colleagues (L'Abbé 1987). This is a plot showing for each study the observed event rate in the experimental group plotted against observed event rate in the control group. L'Abbé plots may be used to view the range of event rates among the trials, to highlight excessive heterogeneity, and, on occasion, to indicate which intervention effect measure may be most consistent across trials. Naïve regression analyses based on L'Abbé plots are misleading, however, since they do not account for sampling error in both observed event rates (Sharp 1996).

L'Abbe plots may be drawn on the scale of the risk (the event rate), the log(risk) or the log(odds) (see Van Houwelingen 1993 for examples of the first and last). At present no advice is available on whether any is preferable in general. The first, however, is most likely to be interpretable by clinicians. An example appears in Figure 5.

Recommendations for L'Abbé plots

7.1. Where interventions are experimental and standard/control, the experimental event rate should be plotted on the vertical axis. When there is no such asymmetry it does not matter which way the plot is done.

7.2. A line indicating no intervention effect should be added.

7.3. Regression lines should not be added (unless they are derived using techniques that account for sampling error in both variables)

7.4. It may be useful to plot points at a size proportional to weight or trial size (preferably weight).

7.5. If the software permits, the graph should be square.

7.6. The raw data (information sufficient to create a 2×2 table from each trial) should be available somewhere in the review.

7.2 Relating intervention effect to 'underlying risk'

A special case of meta-regression is to assess the dependence of intervention effect on control group event rate, on the assumption that the control group event rates reflect the underlying risks of participants in the studies. As Sharp *et al*. explain (Sharp 1996), such regressions may be highly misleading since they can be affected by regression to the mean. Techniques are available that

overcome this problem (Sharp 2000). Simple scatter plots of intervention effect against control group event rate may be misleading, also due to regression to the mean. We recommend that such plots are not presented unless the results of a suitable analysis of the relationship is obtained and superimposed on the plot.

Recommendations for relationship between intervention effect and underlying risk

8.1. Plots should follow recommendations for single variable meta-regression

8.2. The regression line from an analysis specifically designed for underlying risk meta-regression should be superimposed on the plot.

8.3. The raw data (information sufficient to create a 2×2 table from each trial) should be available somewhere in the review.

8 Other graphical displays

In this section we outline two types of graph that have statistical merit but are less familiar to users of Cochrane reviews, and two types of graph in common use but with unproven or poor statistical grounding. These types of graph are not encouraged as part of a Cochrane review, and if used should be accompanied with a sound justification. We close with a brief mention of some other graphs that have been proposed for use within systematic reviews.

8.1 Galbraith (radial) plots

Galbraith has described an alternative to the confidence interval plot for visualising results of studies and meta-analyses (Galbraith 1988, Galbraith1994). His graph has been enthusiastically received by statisticians (Whitehead 1991, Thompson 1993) but may be less readily interpreted by non-statisticians. The plot provides the basis of a simple graphical test for funnel plot asymmetry (Egger 1997). Galbraith plots facilitate examinations of heterogeneity, including detection of outliers.

A Galbraith plot is a plot of a standardized intervention effect (intervention effect divided by its standard error) against the reciprocal of the standard error. Imprecise estimates of effect lie near the origin, and precise estimates further away, giving the correct impression of being more informative. Vertical variation in points describes the extent of heterogeneity. The plot may be interpreted in terms of lines through the origin. Linear regression through the origin of the standardized intervention effects on their inverse standard errors yields a slope equal to the fixed effect meta-analysis estimate. A 'radial' scale (an arc of a circle) allows the determination of any slope, and hence provides details of the unstandardized effect estimates.

Egger *et al*'s test for funnel plot asymmetry is based on the linear regression (*not* confined to passing through the origin) of standardized intervention effects on their inverse standard errors. Statistical significance of the intercept provides a test for funnel plot asymmetry, since under ideal conditions the regression line should pass through the origin.

8.2 Relationship between intervention effect and two or more covariates (meta-regression)

On occasion it may be of interest to investigate the relationship between intervention effect and two or more covariates. Illustration of such a relationship requires three or more dimensions. Lau *et al*. have described the use of response surfaces for the illustration of relationships with two covariates (Lau 1998). Response surface plots and 3-dimensional histograms/bar charts are not encouraged in Cochrane reviews. Two dimensional scatter plots illustrating the relationships between intervention effect and each covariate, and between covariates, may be helpful.

8.3 Survival curves

A standard representation of time-to-event outcomes from clinical trials is a Kaplan Meier curve. These illustrate the survival times of participants in the trial while acknowledging that some were not observed, so that appropriate comparison of the different intervention groups can be made. Kaplan Meier plots from individual trials are suitable for inclusion in Cochrane reviews, though they may easily become too numerous.

Kaplan Meier plots for all pooled participants across trials in a meta-analysis have previously been presented in medical journals. This practice breaks with the principle of comparing like with like. For this reason, until further discussions have taken place the Statistical Methods Group is unable to recommend inclusion of such plots in Cochrane reviews.

8.4 Cumulative meta-analysis

Cumulative meta-analysis (Lau 1995) plots accumulations of studies: this suffers from a lack of independence of points, which could mislead a naïve reader (Antman 1992).

8.5 Further graphical displays

Numerous other graphical displays can sometimes add useful insights to reports of systematic reviews. For example, sequential/prospective meta-analysis (Whitehead 1997, Pogue 1998) may be used to illustrate the accumulation of data with respect to some a priori desirable amount of information. Other suggestions for graphics relevant to meta-analyses include box plots (Light 1994, Petitti 1994), plots related to model diagnostics (Olkin 1995, Hardy 1998), illustrations of distributions (including prior and posterior distributions for Bayesian meta-analyses (Carlin 1992)) and plots to illustrate two-dimensional uncertainty (Thompson 1993, Hardy 1996). Finally, ‘odd-man-out meta-analysis’ (Walker 1988) is a proposal for illustrating summary confidence regions.

8.6 Contributions

This appendix was prepared in 2002-2003 by Julian Higgins on behalf of the Cochrane Statistical Methods Group. The help of the following is particularly appreciated: Doug Altman, Deborah Ashby, Jon Deeks, Gordon Dooley, Diana Elbourne, Sally Hollis, Steff Lewis, Keith O’Rourke, Jonathan Sterne, Simon Thompson and members of the Cochrane Information Management System Group.

9 References

Antman 1992. Antman EM, Lau J, Kupelnick B, Mosteller F, and Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: Treatments for myocardial infarction. *Journal of the American Medical Association* 1992; **268**: 240-248

Arkin 1940. Arkin H and Colton RR. *Graphs: How to Make and Use Them*. New York: Harper and Brothers, 1940

Bailey 1987. Bailey KR. Inter-study differences - how should they influence the interpretation and analysis of results. *Statistics in Medicine* 1987; **6**: 351-360

Carlin 1992. Carlin JB. Meta-analysis for 2 x 2 tables: a Bayesian approach. *Statistics in Medicine* 1992; **11**: 141-158

Cleveland 1994. Cleveland WS. *The Elements of Graphing Data*. Summit, New Jersey: Hobart Press, 1994

Deeks 2001. Deeks JJ, Altman DG and Bradburn MJ. Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In: Egger M, Davey Smith G and Altman DG (eds). *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Books, 2001

Egger 1997a. Egger M, Davey Smith G, and Phillips AN. Meta-analysis: principles and procedures. *British Medical Journal* 1997; **315**: 1533-1537

Egger 1997b. Egger M, Smith GD, Schneider M, and Minder C. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 1997; **315**: 629-634

Galbraith 1988. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Statistics in Medicine* 1988; **7**: 889-894

Galbraith 1994. Galbraith RF. Some applications of radial plots. *Journal of the American Statistical Association* 1994; **89**: 1232-1242

Hardy 1996. Hardy RJ and Thompson SG. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine* 1996; **15**: 619-629

Hardy 1998. Hardy RJ and Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine* 1998; **17**: 841-856

Higgins 2008. Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions* 5.0.0 [updated February 2008]. The Cochrane Collaboration, 2008. Available from www.cochrane.org/resources/handbook.

Hooper 2001. Hooper L, Summerbell CD, Higgins JPT, Thompson RL, Capps NE, Davey Smith G, Riemersma RA, and Ebrahim S. Dietary fat intake and prevention of cardiovascular disease: systematic review. *British Medical Journal* 2001; **322**: 757-763

Johnson 1999. Johnson ES, Lanes SF, Wentworth III CE, Satterfield MH, Abebe BL, and Dicker LW. A metaregression analysis of the dose-response effect of aspirin on stroke. *Archives of Internal Medicine* 1999; **159**: 1248-1253

L'Abbe 1987. L'Abbe KA, Detsky AS, and O'Rourke K. Meta-analysis in clinical research. *Annals of Internal Medicine* 1987; **107**: 224-233

Lau 1995. Lau J, Schmid CH, and Chalmers TC. Cumulative meta-analysis of clinical trials: Builds evidence for exemplary medical care. *Journal of Clinical Epidemiology* 1995; **48**: 45-57

Lau 1998. Lau J, Ioannidis JP, and Schmid CH. Summing up evidence: one answer is not always enough. *The Lancet* 1998; **351**: 123-127

Lewis 2001. Lewis S and Clarke M. Forest plots: trying to see the wood and the trees. *British Medical Journal* 2001; **322**: 1479-1480

Light 1984. Light RJ and Pillemer DB. *Summing Up: The science of Reviewing Research*. Cambridge, Mass: Harvard University Press, 1984

Light 1994. Light RJ, Singer JD and Willett JB. The visual presentation and interpretation of meta-analyses. In: Cooper H and Hedges LV (eds). *The Handbook of Research Synthesis*. New York: Russell Sage, 1994

Olkin 1995. Olkin I. Statistical and theoretical considerations in meta-analysis. *Journal of Clinical Epidemiology* 1995; **48**: 133-146

Petitti 1994 Petitti DB. *Meta-Analysis, Decision Analysis and Cost-Effectiveness Analysis*. New York: Oxford University Press, 1994

Pogue 1998. Pogue J and Yusuf S. Overcoming the limitations of current meta-anlysis of randomised controlled trials. *Lancet* 1998; **351**: 47-52

Schmid 1983. Schmid CF. *Statistical Graphics: Design Principles and Practices*. New York: John Wiley and Sons, 1983

Sharp 1996. Sharp SJ, Thompson SG, and Altman DG. The relation between treatment benefit and underlying risk in metaanalysis. *British Medical Journal* 1996; **313**: 735-738

Sharp 2000. Sharp SJ and Thompson SG. Analysing the relationship between treatment benefit and underlying risk in meta-analysis: comparison and development of approaches. *Statistics in Medicine* 2000; **19**: 3251-3274

Simmonds 1980. Simmonds D (ed). *Charts and Graphs: Guidelines for the Visual Presentation of Statistical Data in the Life Sciences*. Lancaster, UK: MTP Press, 1980

Sterne 2001a. Sterne JAC, Egger M and Davey Smith G. Investigating and dealing with publication bias and other biases. In: Egger M, Davey Smith G and Altman DG (eds). *Systematic Reviews in Health Care: Meta-analysis in Context*. London: BMJ Books, 2001

Sterne 2001b. Sterne JAC and Egger M. Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology* 2001; **54**: 1046-1055

Sutton 1998. Sutton AJ, Abrams KR, Jones DR, Sheldon TA, and Song F. Systematic reviews of trials and other studies. *Health Technology Assessment* 1998; **2**

Thompson 1993. Thompson SG. Controversies in meta-analysis: the case of the trials of serum cholesterol reduction. *Statistical Methods in Medical Research* 1993; **2**: 173-192

Thompson 1994. Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *British Medical Journal* 1994; **309**: 1351-1355

Thompson 2001. Thompson RL, Summerbell CD, Hooper L, Higgins JPT, Little PS, Talbot D, Ebrahim S. Relative efficacy of differential methods of dietary advice: a systematic review. American *Journal of Clinical Nutrition* 2003; **77**: 1052S-1057S

Thompson 2002. Thompson SG and Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 2002; **21**: 1539-1558

van Houwelingen 1993. Van Houwelingen HC, Zwinderman KH, and Stijnen T. A bivariate approach to meta-analysis. *Statistics in Medicine* 1993; **12**: 2273-2284

Walker 1988. Walker AM, Martin-Moreno JM, and Artalejo FR. Odd man out: a graphical approach to meta-analysis. *American Journal of Public Health* 1988; **78**: 961-966

Whitehead 1991. Whitehead A and Whitehead J. A general parametric approach to the meta-analysis of randomised clinical trials. *Statistics in Medicine* 1991; **10**: 1665-1677

Whitehead 1997. Whitehead A. A prospectively planned cumulative meta-analysis applied to a series of concurrent clinical trials. *Statistics in Medicine* 1997; **16**: 2901-2913

Figure 1: Forest plot from a Cochrane review of dietary advice for cholesterol reduction (data from Thompson 2001)

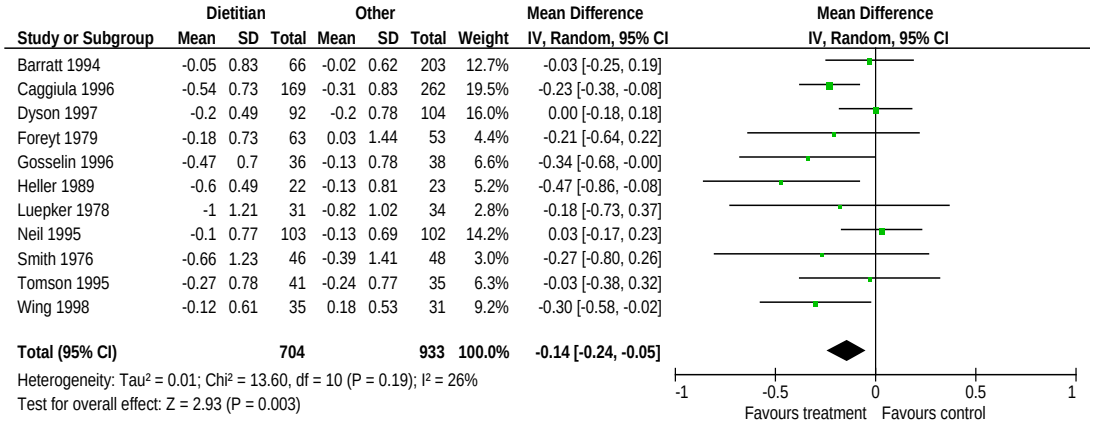

Figure 2: Forest tops plot of subgroup analyses and sensitivity analyses from a review of trials of reduction/modification of dietary fat or cholesterol (data from Hooper 2001)

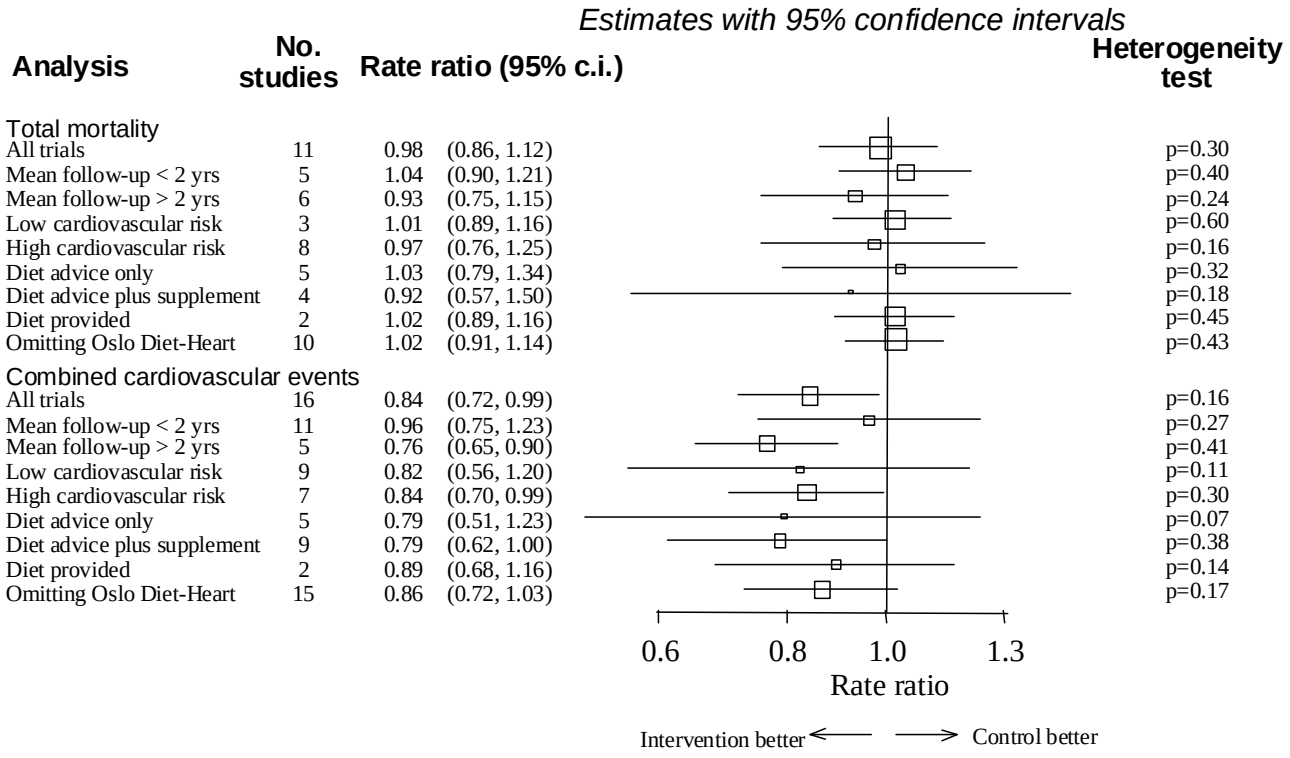

Figure 3: Funnel plot of trials of ACE inhibitors (data from Sterne 2001b)

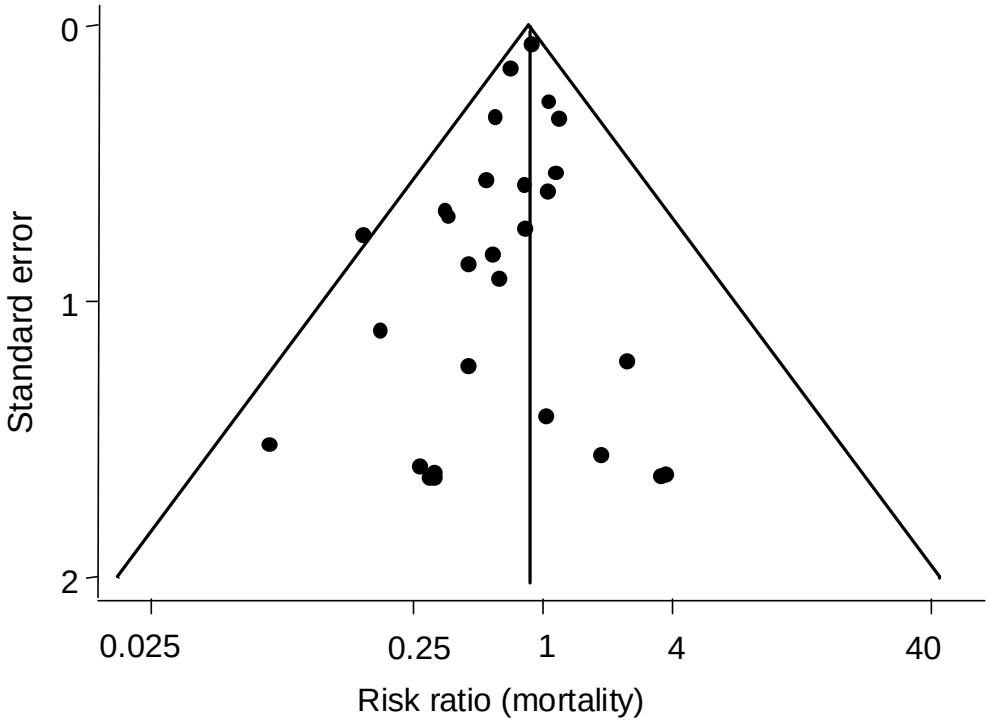

Figure 4: Relationship between relative risk and aspirin dose in 12 trials of aspirin for secondary prevention of stroke (data from Johnson 1999)

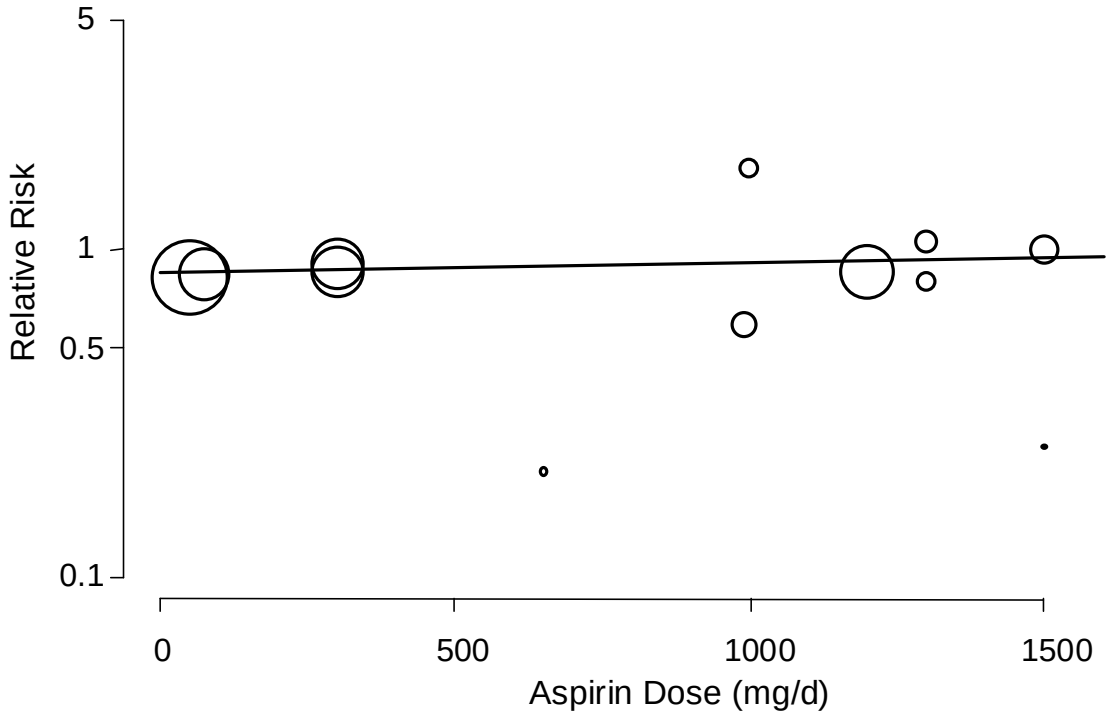

Figure 5: L'Abbé plot of 19 trials of sclerotherapy (data from Sharp 1996)