Research Article

Research Synthesis Methods

Received 9 September 2009, Revised 19 January 2010, Accepted 24 January 2010 Published online 23 March 2010 in Wiley Interscience

(www.interscience.wiley.com) DOI: 10.1002/jrsm.6

Graphical displays for meta-analysis: An overview with suggestions for practice

Judith Anzures-Cabrera[a] and Julian P. T. Higgins[b]*†

Meta-analyses are fundamental tools for collating and synthesizing large amounts of information, and graphical displays have become the principal tool for presenting the results of multiple studies of the same research question. We review standard and proposed graphical displays for presentation of meta-analytic data, and offer our recommendations on how they might be presented to provide the most useful and user-friendly illustrations. We concentrate on graphs that specifically aim to present similar sorts of univariate results from multiple studies. We start with forest plots and funnel plots, and proceed to Galbraith (or radial) plots, L'Abbé (and related) plots, further plots useful for investigating heterogeneity, plots useful for model diagnostics and plots for illustrating likelihoods and Bayesian meta-analyses. **Copyright © 2010 John Wiley & Sons, Ltd.**

Keywords: meta-analysis; graphical displays; forest plot; funnel plot; Galbraith plot; L'Abbe plot

1. Introduction

Based on the widespread understanding that 'one picture is worth ten thousand words' [1], graphical displays are fundamental tools of quantitative research. Graphs enable the visual identification of overall patterns, associations and outlying observations that might be overlooked in tables of data. Although they can be very effective in communicating key aspects of large data sets, if poorly selected or poorly designed they can frustrate and even mislead the reader.

Meta-analyses are fundamental tools for collating and synthesizing large amounts of information, and graphical displays have become the principal tool for presenting the results of multiple studies of the same research question [2]. Here we review standard and proposed graphical displays for presentation of meta-analytic data sets, and offer our recommendations on how they might be presented to provide the most useful and user-friendly illustrations. The paper extends and updates previous reviews of graphical displays for meta-analysis [3–7], and builds on a document previously published online as supplementary material to the *Cochrane Handbook for Systematic Reviews of Interventions* [8]. In collating materials, we supplemented our extensive collection of meta-analysis methods papers with electronic searches of PubMed, Science Citation Index and Social Sciences Citation Index, using search terms such as 'graph*', 'plot*', 'illustrat*' and 'figur*' combined with synonyms for meta-analysis. In order to be able to clarify the quantities being plotted, we overview some basic methods and define some notation in Box 1. However, the mathematical asides can be skipped without hindrance.

We concentrate on graphs that specifically aim to present similar sorts of univariate results from multiple studies. We start by addressing the two principal graphical methods associated with meta-analysis, both available in most univariate meta-analytic situations. These are the forest plot (Section 2) and the funnel plot (Section 3). In the next two sections we discuss Galbraith (or radial) plots and L'Abbé (and related) plots. For each of these first four types of graph we offer our suggestions for implementation. In subsequent sections we review some further plots that are useful for investigating heterogeneity, some plots useful for model diagnostics and some plots for illustrating likelihoods and Bayesian meta-analyses.

[a] *Roche Products Ltd, Welwyn Garden City, U.K.*
[b] *MRC, Biostatistics Unit, Cambridge, U.K.*
* *Correspondence to: Julian P. T. Higgins, MRC, Biostatistics Unit, Cambridge, U.K.*
† *E-mail: julian.higgins@mrc-bsu.cam.ac.uk*

Contract/grant sponsor: MRC; contract/grant number: U.1052.00.011

Copyright © 2010 John Wiley & Sons, Ltd.

Box 1. Notation and basic meta-analysis methods.

A large class of meta-analysis methods is based on the availability of an effect estimate and its standard error from each study. A common, and simple, approach is a weighted average, $M=(\sum w_i y_i)/\sum w_i$, where y_i is the estimate from the study i, and w_i a weight awarded to that study. An approximate standard error for M is given by $1/\sqrt{\sum w_i}$. A fixed-effect meta-analysis, assuming that all y_i are estimating the same underlying effect, can be obtain using weights $w_i^F = 1/s_i^2$, where s_i is the estimated standard error of y_i. A random-effects meta-analysis, allowing the true effects to differ across studies, can be obtained using revised weights $w_i^R = 1/(s_i^2 + \hat{\tau}^2)$, where $\hat{\tau}$ is an estimate of the among-study standard deviation. Note that $\hat{\tau}$ and the within-study standard errors are usually assumed to be known, with the standard errors assumed uncorrelated with the effect estimates. A test for heterogeneity (equivalently of $\hat{\tau}=0$) can be obtained by comparing the statistic $Q=\sum w_i^F (y_i - M^F)^2$ to a χ^2 distribution with degrees of freedom equal to one less than the number of studies (with M^F the meta-analysis point estimate with fixed-effect weights). This statistic can also be used as the basis for $\hat{\tau}$ [9].

In many meta-analyses, the estimate y_i is obtained as a difference between independent summary statistics from two groups. For example, it may be the difference between two means (in which case y_i is a mean difference), the difference between two rates (y_i is a rate difference), or the difference in logit probabilities (y_i is a log odds ratio). For such circumstances, we write, $y_i = x_{1i} - x_{2i}$ with x_{1i} and x_{2i} being the summary statistics from group 1 and group 2, respectively. Note that standardized mean differences do not fall conveniently into this category because the standardizing standard deviation is computed from both groups. Furthermore, results from matched studies and single group studies are not constructed in this way.

2. Forest plots

2.1. Overview

Forest plots, also known as confidence interval plots (but not properly as 'forrest plots' [10]), are probably the most familiar method for presenting results of meta-analyses. A forest plot displays effect estimates and their confidence intervals for each study and, usually, the meta-analysis [11–13]; a brief history can be found in Lewis and Clarke [10] and a study of their use in practice in Schriger *et al.* [14]. Effect estimates may be virtually any quantity that can be estimated from every study. Typical examples are odds ratios (or log odds ratios), standardized mean differences (often called 'effect sizes'), correlations (of transformations thereof), proportions (or transformations thereof), risk ratios (or log risk ratios), hazard ratios (or log hazard ratios), means, and differences in means. For effect estimates with sampling distributions that are not well approximated by a normal distribution, it is usual for the forest plot to be specified on a transformed scale (e.g. log risk ratio), with axis labels presented on the more intuitive scale (e.g. risk ratio) [15]. An example of a meta-analysis of odds ratios is given in Figure 1, taken from a Cochrane review on the effects of antibiotics to treat otitis media in children [16].

Mathematically, forest plots typically illustrate the effect estimates, y_i, for each study along with a confidence interval, usually constructed as $y_i \pm Z^{\alpha} \times s_i$ to create a two-sided $100(1-\alpha)\%$ confidence interval, where Z^{α} is the $100(1-\alpha/2)^{th}$ percentile of a standard normal distribution. It has become standard to use the standard normal distribution to compute confidence intervals irrespective of the nature of the effect estimates, and this corresponds to the usual assumption that the within-study standard errors are known (Box 1). However, for certain effect measures, such as mean differences, it might be preferable to use an appropriate t-distribution.

In Figure 1, each study is represented by a square at a point estimate of effect and a horizontal line extending either side of the block to depict a 95% confidence interval. The area of the block is proportional to the weight assigned to that study in the meta-analysis. The confidence interval and the area of the block convey similar information. However, the use of different block sizes is a device to draw the eye towards the studies with larger weight (smaller confidence intervals). Failure to use this device may result in unnecessary attention being attracted to those smaller studies with wider confidence intervals that put more ink on the page (or more pixels on the screen). Studies may be divided into subsets for presentation in a forest plot, for example by types of participants (as in Figure 1), by time points or by assessments of study validity.

Forest plots may include the result of the overall effect from a meta-analysis, normally at the bottom of the graph, and often using a diamond to distinguish it from the individual studies. It is common to plot the effect measure on the horizontal axis, in contrast to the usual convention of plotting dependent variables on the vertical axis. This has the advantage of allowing study identifiers and detailed data to be plotted alongside the results. Some authors prefer to plot the effect measure on the vertical axis [17]. No general recommendation is appropriate for the order in which the studies should be presented in a forest plot. Many authors default to an alphabetical ordering for ease of cross-referencing with tables and reference lists. Other useful orders include a chronological ordering [18] or an ordering by weight (or precision), which creates an approximate version of a funnel plot (see Section 3).

Copyright © 2010 John Wiley & Sons, Ltd.

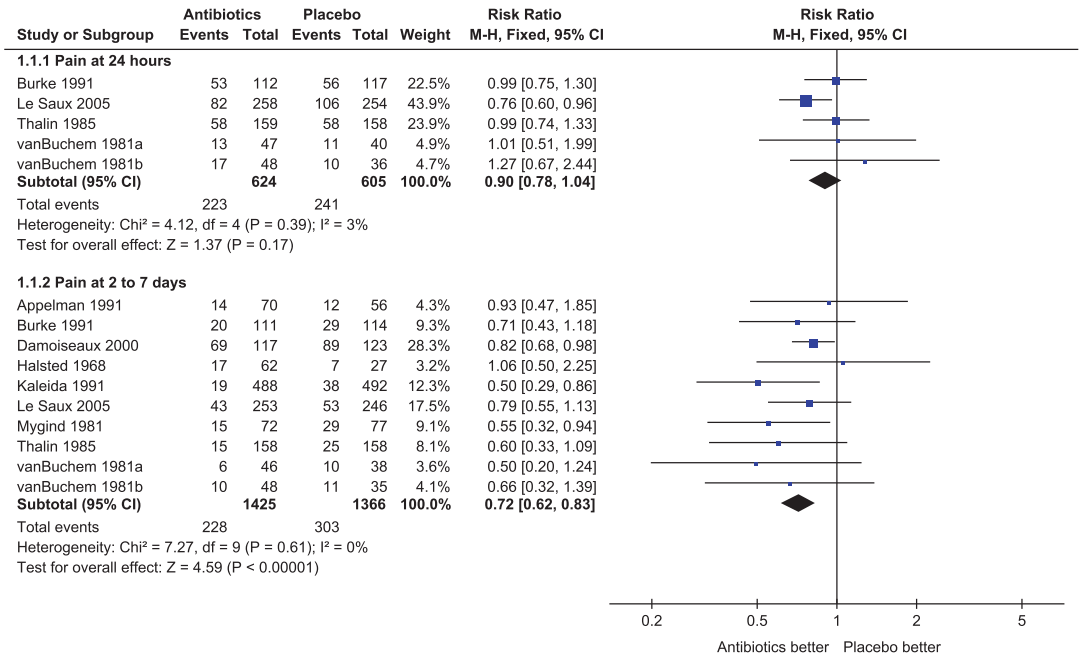

Figure 1. Forest plot. Antibiotics for acute otitis media in children: Effect on recovery from pain [16]. Copyright Cochrane Collaboration, reproduced with permission.

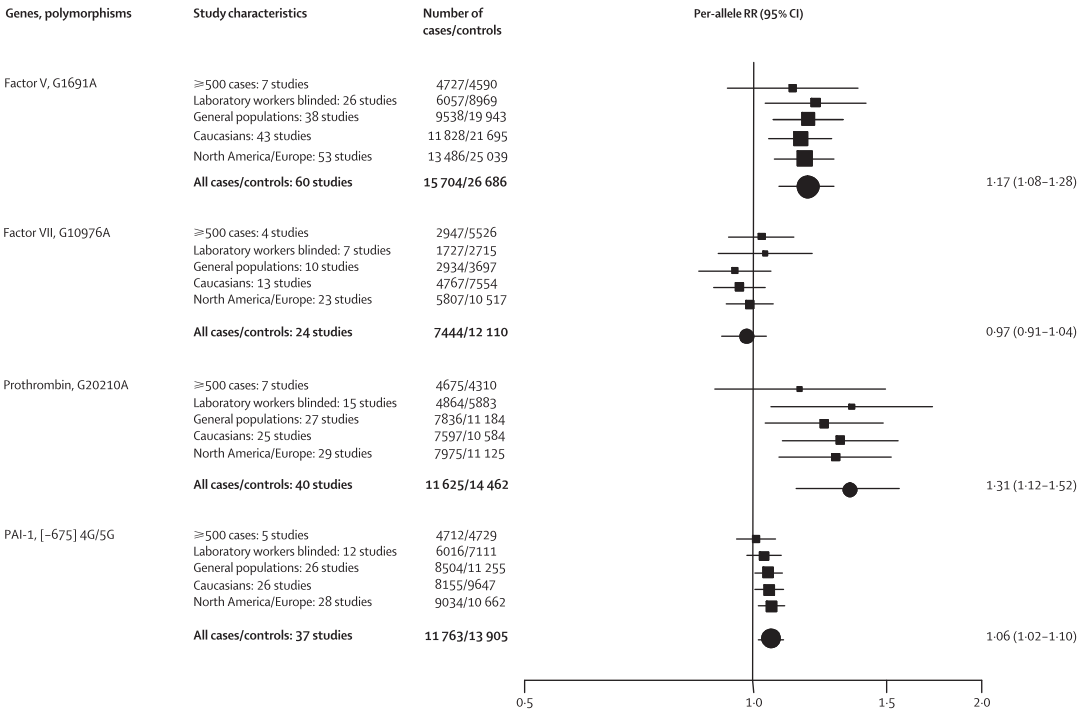

Figure 2. Summary forest plot. Meta-analysis of studies of coronary diseases and four haemostatic gene polymorphisms, grouped by various characteristics [19]. Reproduced with permission.

2.2. Variations on forest plots

Forest plots may 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. This is a particularly useful form of graph, and we use the name 'summary forest plot' to indicate that the individual points represent meta-analyses rather than studies. Figure 2 provides an example of a summary forest plot illustrating a series of subgroup analyses in a review of association between variants of certain haemostatic genes and coronary heart disease [19].

Copyright © 2010 John Wiley & Sons, Ltd.

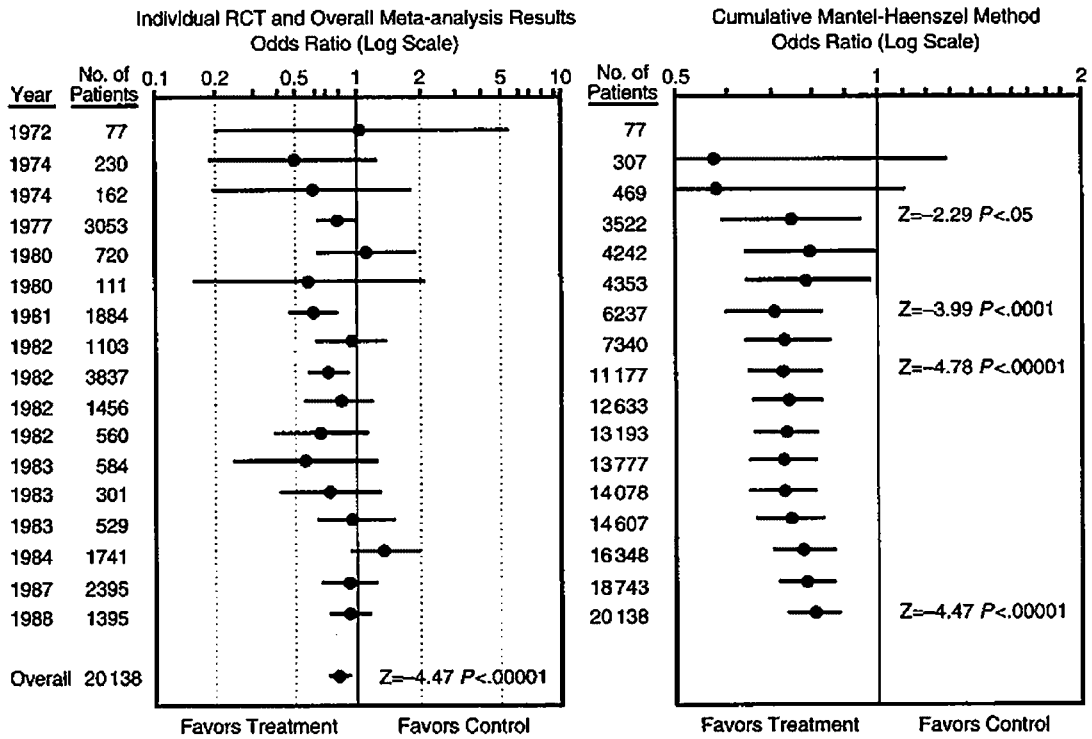

Figure 3. Cumulative meta-analysis. Results of 17 randomized control trials of the effects of oral β-blockers for secondary prevention of mortality in patients surviving a myocardial infarction presented as two types of meta-analyses [25]. On the left is the traditional one, revealing many trials with non-significant results but a highly significant estimate of the pooled results. On the right, the same data are presented as cumulative meta-analyses, illustrating that the updated pooled estimate became statistically significant in 1977 and has remained so up to the present. Note that the scale is changed on the right graph to improve clarity of the confidence intervals. Copyright © 1992 American Medical Association. All rights reserved.

Whether or not the individual study results are present, a forest plot can usefully illustrate the results of different methods for meta-analysis. Perhaps the most obvious example of this is the inclusion of both fixed-effect and random-effects meta-analysis results. For Bayesian or empirical Bayes meta-analyses, it is conventional to include shrunken estimates and credibility intervals from the individual studies below the usual observed result [20–22]. Prediction intervals may be added to summary results from random-effects meta-analyses (classical or Bayesian) to illustrate heterogeneity of effects, which is not conveyed by the confidence interval around the mean effect [23].

Two further variations of summary forest plots are cumulative meta-analysis plots and 'leave-one-out' sensitivity analysis plots. Cumulative meta-analyses show the results of a series of meta-analyses, one performed after the addition of each study in sequence. When using the chronological sequence of studies, it is easy to identify the point in time at which, for example, there was evidence that a medical intervention was beneficial or harmful [24]. A classic example of a cumulative meta-analysis plot appears in Figure 3, from a paper that compares accumulated evidence with recommendations from clinical experts in the treatment of myocardial infarction [25]. Lau and colleagues [26] note that a variety of study characteristics can be used to order clinical trials in a cumulative meta-analysis, including date of publication, control group risk, study size, magnitude of treatment effect and quality of individual trials. However, we prefer meta-regression approaches to explore the impact of most of these characteristics, and consider cumulative meta-analyses to have their primary appeal in depicting what would have happened had a meta-analysis been performed at each in a series of time-points, and not as a means of drawing statistical inferences. Successive results in a cumulative meta-analysis are not independent, and could mislead the naïve reader [25]. Pogue and Yusuf [27] suggest using a stopping boundary in cumulative meta-analysis plots to determine the point in time in which convincing conclusions for the evidence of the effect of a treatment can be undertaken. Within a similar context, Ioannidis *et al*. [28] use an inverted cumulative meta-analysis graph to present the results of a 'recursive cumulative meta-analysis', which can be seen as series of cumulative meta-analyses.

A useful sensitivity analysis is one in which the meta-analysis is repeated, each time omitting one of the studies. A plot of the results of these meta-analyses, called an 'exclusion sensitivity plot' by Bax *et al*. [29], will reveal any studies that have a particularly large influence on the result of the meta-analysis.

Finally, we note that simple lines depicting confidence intervals may suggest that all values in the interval are equally likely. This is not the case: those towards the middle are more plausible than those towards the ends. This can be expressed statistically using the notion of likelihood, and Barrowman and Myers propose to illustrate the likelihoods directly within the forest plot [30] by making lines for both studies and meta-analyses widest at the point estimate (the maximum likelihood estimate), tapering towards the ends in a way that is directly proportional to the likelihood, reaching a point at the limits of a 95% likelihood-based confidence interval. In their own words, 'we introduce a new kind of display, the *raindrop plot*, so called because the visual effect is reminiscent of raindrops, streaking across a car window'.

Copyright © 2010 John Wiley & Sons, Ltd.

2.3. Our suggestions

We offer the following suggestions to ensure that forest plots are informative and clear.

- Forest plots should be included whenever feasible and appropriate in the report of a meta-analysis. When studies are too numerous to present individually, summary forest plots should be considered.
- It is usually preferable to plot the effect measure on the horizontal axis, so that rows (depicting studies or meta-analyses) can be annotated. Indeed, it should always be possible to determine what is represented by each entry in the plot (i.e. which study or which analysis).
- Ratio measures of effect (such as odds ratios, relative risks, hazard ratios and rate ratios) should be plotted on the log scale. The tick marks and labelled values on the axis, however, should be on the original (anti-logged) scale [15]. Similar strategies should be followed for other transformed measures (e.g. the log or logit transformation of a probability, or the Fisher Z transform of correlation coefficient).
- A reference line should be drawn at the position of no effect, if such a null value is relevant. It may be helpful to add another line (e.g. a dashed line) to indicate the estimated summary effect.
- 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 the summary effect, though any differences should be clearly labelled.
- The size of the block representing a point estimate from a study should usually relate to the amount of information in the study. This is typically achieved by making its area proportional to its inverse variance. 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).
- Where directions of effect have an interpretation that could be unclear, it is important to provide an explanation. A useful device is to include indications directly below the plot (for example, 'Favours group therapy ←' and '→ Favours individual therapy').
- All data being represented in forest plots (or results sufficient to calculate them) should be presented numerically (for example, on the forest plot, in a table, or in web-based supplementary materials).

3. Funnel plots

3.1. Overview

Funnel plots, introduced by Light and Pillemer [31] and discussed in detail by Sterne *et al.* [32] and Egger *et al.* [33], are useful precursors or adjuncts to meta-analyses. A funnel plot is a scatter plot of effect estimate against a measure of precision (or study size). Mathematically, it plots y_i against some function of s_i. An example of a funnel plot appears in Figure 4, in which the effect estimates are plotted against their standard errors for studies of association between body mass index and pre-menopausal breast cancer [34].

Funnel plots are used primarily as a visual aid for detecting bias or heterogeneity, and often it is not possible to distinguish between the two. A complete collection of unbiased studies, all estimating the same underlying effect with different levels of

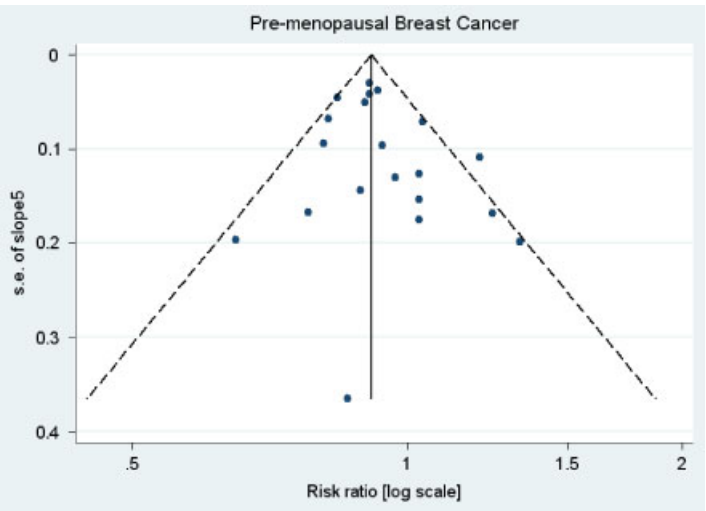

Figure 4. Funnel plot. Funnel plot for a meta-analysis of association between body mass index and pre-menopausal breast cancer [34]. Reprinted with permission from Elsevier.

Copyright © 2010 John Wiley & Sons, Ltd.

precision, would produce a symmetric funnel shape, with increasing scatter with decreasing precision. An asymmetric funnel indicates a relationship between effect size and precision in the studies at hand. This is most likely to be due to reporting bias, to a systematic difference between smaller and larger studies, or to the presence of subsets of studies with different mean effect sizes. Reporting bias in this context refers to the non-availability of results from small studies with non-significant results, and may be due to suppression of the whole study or suppression of data relevant to the meta-analysis at hand. Systematic differences between smaller and larger studies can be caused by differential study validity in smaller compared with larger studies, or genuine differences in effects in studies of different sizes. If there are subsets of studies with different means, then an asymmetric appearance may be due to the presence of several symmetric funnels. Use of different plotting symbols for studies with different characteristics can help to distinguish among these possibilities [18, 35]. Asymmetry can also arise, or be exaggerated, with an inappropriate choice of effect measure, with an inappropriate choice of precision measure, with multiple inclusion of smaller (or larger) studies, or by chance [33, 36, 37]. Whatever the cause, an asymmetric funnel plot leads to doubts over the appropriateness of a simple meta-analysis and suggests the need to investigate possible causes.

There are several possibilities for the choice of measure of precision in a funnel plot, including total sample size, standard error, inverse variance (weight) or inverse standard error. Sterne and Egger [38] compare these measures with others, and recommend the standard error. When the standard error is used, straight lines may be drawn to define a region within which 95% of points would be expected to lie in the absence of both heterogeneity and publication bias.

Like forest plots, funnel plots are often drawn with the effect measure on the horizontal axis, with precision on the vertical axis, making it easier to assess symmetry visually. However, eye-balling funnel plots can be difficult, and studies have shown that it may not reliably distinguish between plots of similar shape with and without publication bias [39].

3.2. Variations on funnel plots

To help assess the symmetry of a funnel plot, Elvik proposes to display the mode and median effect estimates as an indication of skew [18]. Peters *et al.* propose adding contours to a funnel plot that represent statistical significance of individual study findings [40]. This aids the interpretation of any funnel plot asymmetry that is identified, since if 'holes' in the plot correspond to studies that would have been statistically significant, this may cast doubt on a conclusion that publication bias is operating. Ferrer suggests that a plot of effect size against year of publication (rather than precision) might sometimes be expected to appear funnel-shaped [4]; this will be the case if studies become either more consistent or more precise over time.

Some methods for identifying or addressing potential publication bias are based on the funnel plot. Regression tests focus on the slope of a straight-line fit to points on a funnel plot using standard error as a measure of precision [33, 41], and the trim-and-fill method 'trims' studies from one side of the plot and 'fills' them in on both sides of the plot, in an attempt to deduce the results from both available and unavailable studies under a particular type of publication bias [42]. A funnel plot including the original studies and the imputed, hypothetical unpublished, studies (using a different plotting symbol) is a useful way to illustrate the results of a trim-and-fill analysis [37].

3.3. Our suggestions

We offer the following suggestions to ensure that funnel plots are informative and clear.

- Funnel plots with fewer than 10 studies should be avoided or interpreted with great care [43].
- We consider funnel plots to be easier to interpret when the effect measure is plotted along the horizontal axis.
- Ratio measures of intervention effect (such as odds ratios, relative risks, hazard ratios and rate ratios) should be plotted on the log scale. The tick marks and labelled values on the axis, however, should be on the original (anti-logged) scale. Similar strategies should be followed for other transformed measures (e.g. the log or logit transformation of a probability, or the Fisher Z transform of correlation coefficient).
- The measure of study size (on the vertical axis) should generally be the standard error of the 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 [38].
- Points should all be the same size, since the size of a study is already described using the vertical axis.
- It is helpful to include 95% limit lines based on a fixed-effect meta-analysis, to help indicate the extent of heterogeneity.
- We consider contours to be a particularly useful addition to help assess the likelihood that asymmetry is due to publication bias [40].
- All data being represented in funnel plot (or results sufficient to calculate them) should be presented numerically (for example, in a table, or in web-based supplementary materials), and it should be possible to determine (usually through this mechanism) which point relates to which study.

4. Galbraith (radial) plots

Galbraith proposes an alternative to the forest plot for visualizing results of studies and meta-analyses [15, 44]. His graph has been enthusiastically received by statisticians [45–47], but might be less readily interpreted by non-statisticians. Galbraith plots facilitate examinations of heterogeneity, including detection of outliers.

Copyright © 2010 John Wiley & Sons, Ltd.

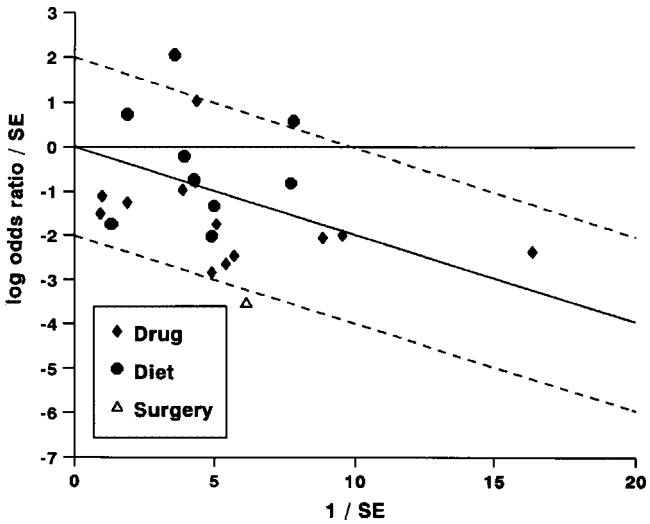

Figure 5. Galbraith plot. Galbraith plot of the log odds ratio for ischaemic heart disease in trials of serum cholesterol reduction by type of intervention [46]. Copyright Sage Publications. Reproduced with permission.

Mathematically, a Galbraith plot is a scatter plot of standardized effect estimates (estimates divided by their standard errors, or y_i/s_i) against inverse standard error ($1/s_i$). Imprecise estimates of effect lie near the origin and precise estimates further away, so that distance from the origin conveys the relative amounts of information in the different studies. An unweighted regression line constrained through the origin has a slope equal to $(\sum(1/s_i)(y_i/s_i))/\sum(1/s_i)^2 = (\sum w_i^F y_i)/\sum w_i^F$, that is the fixed-effect meta-analysis point estimate. Similarly, lines from the origin to each individual point have slope equal to the (unstandardized) point estimate in each study. A radial (arc of a circle) scale can be used to illustrate these slopes, allowing immediate derivation of point estimates from the studies and the meta-analysis. When the radial scale is added, Galbraith plots are sometimes called 'radial plots'. A statistical test for non-zero intercept in an unconstrained line through the plotted points is equivalent to a particular test for funnel plot asymmetry [33].

Vertical scatter of points in a Galbraith plot reflects the extent of heterogeneity. Lines drawn at a vertical distance of ± 2 above and below the regression-through-the-origin line represent an approximate 95% confidence region (Figure 5). Under a fixed-effect meta-analysis model, 95% of studies will, on average, lie between these two lines, and an experimental study has shown that assessors of Galbraith plots can successfully diagnose simulated heterogeneity when it is present [6]. Subsets of studies can be differentiated through the use of different symbols, which may help to identify possible sources of heterogeneity.

4.1. Our suggestions

We offer the following suggestions to ensure that Galbraith plots are informative and clear.

- Galbraith plots might be particularly appropriate when there are more studies than can comfortably be displayed on a forest plot.
- Points should not be drawn in sizes proportional to the weights assigned to the studies in the meta-analysis, since this information is already represented by the horizontal axis.
- Parallel lines representing confidence intervals for the summary effect are useful.
- The horizontal and vertical axes in a Galbraith plot are not usually of substantive interest. Instead, a radial axis allows the results of each study to be determined (perhaps with the aid of a ruler). The tick marks and labelled values on the radial axis might be presented on the original scale of the metric (e.g. odds ratio rather than log odds ratio).
- Where directions of effect have an interpretation that could be unclear, it is important to provide an explanation. A useful device is to include indications above and below the horizontal axis (for example, 'Favours group therapy' and 'Favours individual therapy').
- All data being represented in the Galbraith plot (or results sufficient to calculate them) should be presented numerically (for example, in a table, or in web-based supplementary materials), and it should be possible to determine (usually through this mechanism) which point relates to which study.

5. L'Abbé plots and related plots

5.1. Overview

This section relates to plots that are relevant only to meta-analyses of studies comparing two groups, such as experimental studies of two interventions (including randomized trials), or studies of diagnostic test accuracy in which sensitivity and specificity are jointly of interest. The L'Abbé plot was introduced in 1987 in the context of meta-analyses of clinical trials with dichotomous

Copyright © 2010 John Wiley & Sons, Ltd.

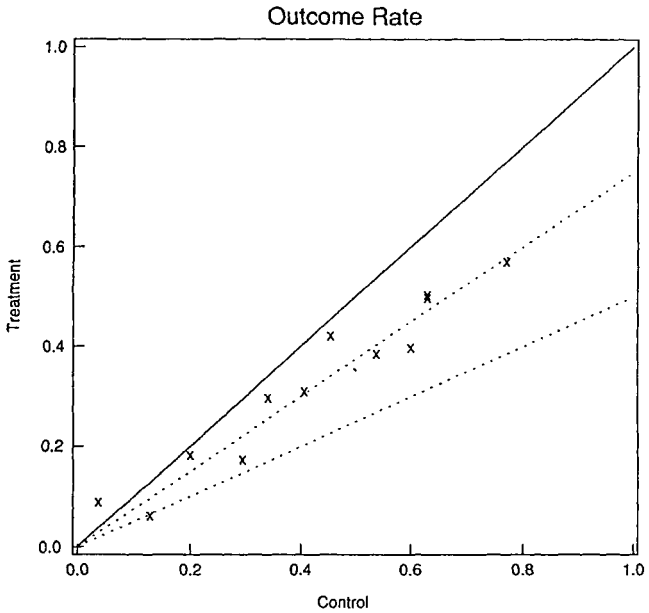

Figure 6. L'Abbé plot. Study outcomes derived from a computer simulation designed to show consistent (homogeneous) treatment effects, measured by proportionate risk reductions [48]. Copyright American College of Physicians. Reproduced with permission.

outcomes, as a plot of observed risks in the treatment group against observed risks in the control group [48]. L'Abbé plots allow inspection of the range of risks among the trials, to highlight excessive heterogeneity, and to indicate which treatment effect measure may be most consistent across trials. L'Abbé plots may be drawn on the scale of the risk, the log risk or the log odds (see Van Houwelingen *et al*. for examples of the first and last [49]). In the notation of Box 1, the L'Abbé plot is of x_{1i} against x_{2i}, where these could represent risks, log risks or logit risks for the two groups. An example is provided in Figure 6.

Since no information about study precision is included in the observed summary data plotted in a L'Abbé plot, it is advisable to use plotting symbols proportional to the precision of the study estimate. Some authors use different symbols according to study-level covariates as a means of investigating heterogeneity.

5.2. Variations on L'Abbé plots

The principle of the L'Abbé plot can be applied to a meta-analysis of any type of study that compares two independent groups. For example, in clinical trials with continuous outcomes, treatment group means may be plotted against control group means [50].

Results of meta-analyses may be overlaid on L'Abbé plots. Van Houwelingen *et al*. present results of a bivariate meta-analysis, including an ellipse that contains a 95% confidence region for the estimated parameters of the random-effects distribution, the empirical Bayes estimates for each study as well as the overall meta-analysis result [21, 49]. Jimenez *et al*. superimpose contours onto the L'Abbé plot for constant treatment effect, using different plots for different metrics (risk difference, risk ratio, odds ratio) [51]. By illustrating the meta-analysis estimate and its confidence interval on each of these different metrics, they argue that such a graph is useful to determine which scale is the most appropriate. Deeks prefers to plot contour lines of the absolute benefit of treatment against baseline (control group) risk [52]. He expresses absolute benefit as cases prevented per 100 treated for a given baseline risk, calculated according to the particular effect metric. In addition to the contour lines, his graphs include shaded areas for absolute benefits that are unachievable.

For studies of diagnostic or screening test accuracy, the same construction as the L'Abbé plot is often used to summarize study results and meta-analyses. Rather than plotting the proportion of events among the exposed against proportion of events among the unexposed, a receiver operating characteristic (ROC) plot illustrates proportion of test-positives among true-positives (sensitivity) against proportion of test-positives among true-negatives (1 – sensitivity). When applied to a single study, each point indicates how the test performs as the threshold varies for determining a positive versus a negative result in the test, and points corresponding to different thresholds fall on the ROC curve. In meta-analysis, points are plotted for the different studies [53]. If different thresholds were the only differences between studies, these points would lie on the ROC curve. In practice, this will not be the case, and the scatter of points from different studies gives an idea of variability among studies. A summary ROC (SROC) curve may be fitted to the points, representing a summary estimate of test accuracy [54].

5.3. Our suggestions

We offer the following suggestions to ensure that L'Abbé plots and similar plots are informative and clear.

- L'Abbé plots might be considered whenever each study contributes two statistically independent pieces of information that might be expected to be inherently correlated across studies (such as treatment and control group outcomes; or sensitivity and specificity).

Copyright © 2010 John Wiley & Sons, Ltd.

- The graph should be square in shape. Where there is an experimental group and a control group, the experimental group should be plotted on the vertical axis.
- A (diagonal) line indicating no effect is often helpful.
- It may be useful to plot points at a size proportional to weight or trial size (preferably weight), since this information is not otherwise conveyed.
- A meta-analysis result can be included, and the plot forms a particularly convenient way to illustrate the results of bivariate approaches to meta-analysis.
- All data being represented in the L'Abbé plot (or results sufficient to calculate them) should be presented numerically (for example, in a table, or in web-based supplementary materials), and it should be possible to determine (usually through this mechanism) which point relates to which study.

6. Further plots for detecting and describing heterogeneity among studies

6.1. Introduction

Heterogeneity in meta-analysis refers to differences in underlying effects, so that estimates are more variable across studies than would be expected by chance alone. Statistical tools are available to assess the extent of heterogeneity, such as the χ^2 test mentioned in Box 1, estimation of the among-study standard deviation (τ) and, more recently, computation of the I^2 statistic [55]. Forest plots, funnel plots, Galbraith plots and L'Abbé plots can all be used informally to identify heterogeneity, which manifests itself in poor overlap of confidence intervals in a forest plot, by scatter beyond 95% confidence bounds in a funnel plot, by scatter beyond the ± 2 lines in a Galbraith plot and by a general dispersion of points in a L'Abbé plot. The forest plot has been observed to be the most successful of these at diagnosing simulated heterogeneity, with high reproducibility [6]. A number of further graphical methods can assist in the identification or investigation of heterogeneity, and we review the principal proposals below.

6.2. Simple summaries of effect estimates

Histograms, box plots, stem-and-leaf plots and other representations of the distribution of effect estimates across studies may reveal notable departures from the assumptions of standard meta-analytic models [3, 18]. However, these do not account for uncertainty in the estimates, and apparent skew may be unimportant when precision is taken into account. We therefore do not advocate the use of such plots.

6.3. Odd man out

Walker *et al*. propose a graphical approach they call 'odd man out' [56]. This is based on a forest plot, in which portions of the effect scale are shaded to produce a summary confidence region. The region represents those effects that are included in at least $(N-1)$95% confidence intervals out of N studies. For moderate numbers of studies (five to ten), this interval has approximately 95% coverage under a fixed-effect meta-analysis model. If there is heterogeneity among the studies, the graph may show two disjoint areas. Several authors have implemented this approach [57, 58].

6.4. Normal probability plots

Several authors have proposed the use of normal probability plots to investigate the nature of heterogeneity. Hardy and Thompson [59] plot the contribution of each study to the test statistic for heterogeneity against quantiles of the standard normal distribution, as well as a similar plot based on a random-effects meta-analysis. Mathematically, the contribution from study i to the Q statistic is $q_i^F = \sqrt{w_i^F}(y_i - M^F)$; for a random-effects model, $q_i^R = \sqrt{w_i^R}(y_i - M^R)$ is plotted instead. Under a true fixed-effect model, either plot should produce a straight line through the origin with slope equal to 1 (providing studies are sufficiently large). Under a true random-effects model, the normal plot based on q_i^F will produce a straight line with slope greater than 1, whereas that based on q_i^R should again have a slope equal to 1. Departures from these scenarios may indicate forms of heterogeneity that require further investigation. Wang and Bushman plot standardized effect estimates from the studies (mathematically, $q_i = y_i / s_i$, equivalently $q_i = \sqrt{w_i^F} y_i$), and use this to assess the normality assumption (as above), to investigate whether all studies appear to come from different populations (in which case an S-shaped curve might be observed), and to look for publication bias (in which case standardized effect estimates near to zero may be missing, resulting in either a 'step' or a U-shaped curve) [60]. Confidence interval lines may be added to the plot to help indicate departure from the normality assumption [37, 60]. Normal plots have also been used to test assumptions of normality for Bayesian meta-analyses [17, 22, 61]. An alternative to a normal probability plot is simply to draw a histogram of standardized (q_i) or normalized (q_i^F or q_i^R) scores. These should resemble a normal distribution if the assumptions of the model hold [62]. Histograms of the contributions to the heterogeneity statistic (q_i^F) have been found to be associated with successful and reproducible identification of simulated heterogeneity [6].

Copyright © 2010 John Wiley & Sons, Ltd.

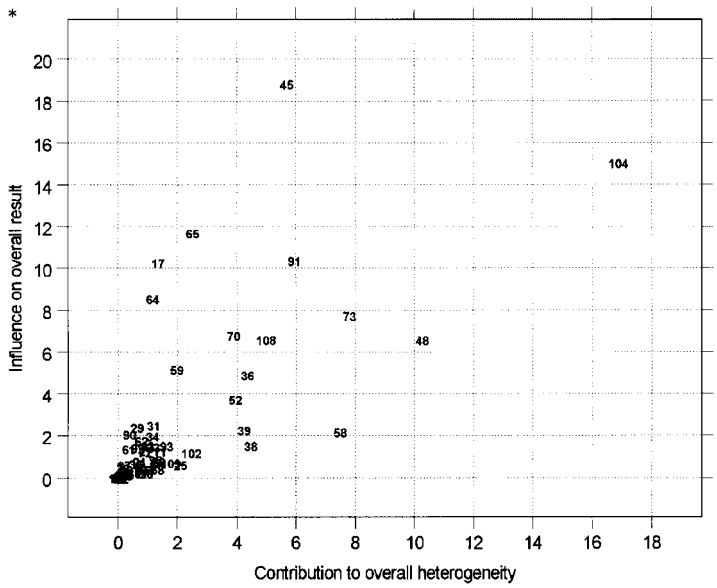

Figure 7. Detecting heterogeneity. Baujat plot for a meta-analysis of chemotherapy in head and neck cancer [63]. Copyright John Wiley & Sons Limited. Reproduced with permission.

6.5. Baujat plot

Baujat *et al*. propose a graphical method to identify studies that impact on heterogeneity or on the meta-analytic estimate, or both [63]. Like Hardy and Thompson, they plot the contribution to the heterogeneity statistic for each study (q_i^F) on the horizontal axis. Along the vertical axis they plot the influence of the study on the overall effect, measured as the squared difference between meta-analysis estimates with and without the *i*th study, divided by the estimated variance of the meta-analysis estimate without the *i*th study. The rationale is that exclusion of an outlying, influential study will affect the meta-analytic estimate, and the effect will be more exaggerated in the plot if the excluded study contributes a relatively large proportion of the weight to the full meta-analysis. Studies located in the upper right corner of the graph will be the most influential ones with the highest contribution to the overall heterogeneity, as illustrated in Figure 7.

7. Plots for describing the relationship between effect estimates and covariates (meta-regression)

7.1. Single covariate

Often a source of heterogeneity can be summarized as a study-level covariate, i.e. some varying characteristic of the studies. Meta-regression is the statistical analysis of the association between effect size and the value of one or more study-level covariates. Potential candidates include measures of quality or validity, study location, intervention characteristics and population characteristics. A simple scatter plot with the covariate along the horizontal axis and the treatment effect along the vertical axis provides a convenient visual impression of the relationship. It is important to reflect the precision of the effect estimates in the plot. Often this is achieved by sizing the plotting symbols (for example, open circles) to reflect the precision of each effect estimate [64]. The areas of these symbols may be based either on fixed-effect weights (i.e. within-study sampling error) or random-effects weights. Alternatively, 95% confidence intervals can be included for each effect estimate [65]. Bailar uses equally sized points without confidence intervals, but indicates the size of each study next to its point [66]. A limitation of many scatter plots is that individual studies cannot easily be identified. To overcome this, Berkey *et al*. display numbers rather than points in a scatter plot [67], while Bailar annotates each point [66]. The meta-regression analysis yields a line that may be superimposed on the scatter plot. An example from a meta-analysis of the effect of electronic fetal heart rate monitoring on perinatal mortality is provided in Figure 8, in which the reduction in risk associated with the intervention is plotted against study publication date as a covariate.

7.2. Multiple covariates

On occasions it may be of interest to investigate the relationship between treatment effect and two or more covariates. Illustration of such a relationship requires three or more dimensions. Lau *et al*. describe the use of response surfaces for the illustration of relationships with two covariates [69]. When second or subsequent covariates may be collapsed into categories, a simpler approach is exemplified by Lewington *et al*. [70]. They plot the effect estimate against the first covariate (for example blood pressure) and separate the effect by subgroups given by further covariates (for example, by age groups). The result is a graph

Copyright © 2010 John Wiley & Sons, Ltd.

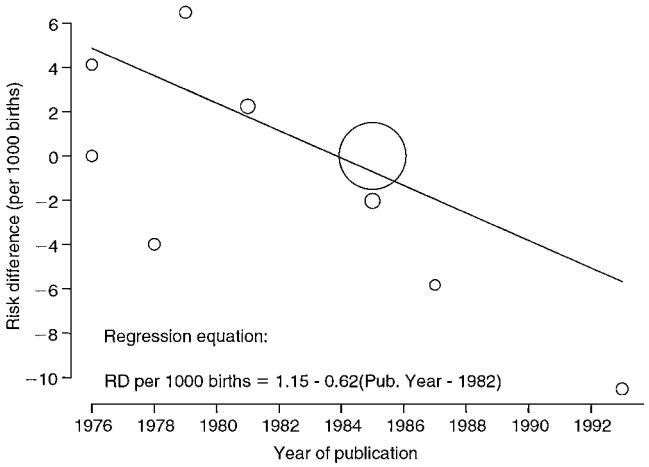

Figure 8. Meta-regression plot. Meta-regression model including a random effect and using the covariate year of publication to assess the use of electronic fetal heart rate monitoring in perinatal mortality [68]. Copyright Sage Publications. Reproduced with permission.

that presents meta-regression lines for different age groups. Non-parallel lines would indicate the presence of interaction between the covariates.

8. Plots useful for sensitivity analysis and diagnostics

8.1. Introduction

Many assumptions are made in the computation of meta-analytic results, and these should be assessed wherever practicable. Sensitivity analyses determine whether the assumptions have an important impact on the results and conclusions of the meta-analysis, while diagnostic procedures can evaluate the suitability of a particular modelling assumption to the data. Plots can be invaluable for these processes; indeed Olkin writes, 'Plot, plot, plot whenever and whatever you can.' [71]. We have already indicated the use of certain plots for sensitivity analysis, such as forest plots for illustrating various models and methods or for the impact of excluding each study in turn, and 'trim-and-fill' results in funnel plots for the potential impact of missing studies on the meta-analysis. The normal plot described in Section 6 can be used in diagnosing the suitability of a normal distribution for random effects; histograms and box plots might also reveal non-normality [6], although as we note above we do not recommend these since they do not account for within-study precision.

8.2. Dependence of results on unknown quantities

One further category of plots can produce illuminating results and is in our view under-used in meta-analysis. These are plots that systematically illustrate the way in which key findings depend on unknown quantities or other factors. We provide three examples of such plots. First, it is well known that the heterogeneity variance in a random-effects meta-analysis is imprecisely estimated when the number of studies is not large [23]. Thompson is representative of several authors who have plotted the meta-analysis estimate (M) as a function of the among-study variance (τ^2) to show the sensitivity of the main result to the among-study variance [46]. Such plots can illustrate the full range of possibilities from zero variance (corresponding to a fixed-effect analysis) through a standard random-effects variance estimate to infinite variance (corresponding to equal weighting of studies). If the summary effect does not change much for different values of the among-study variance, then its estimation is robust to this common source of uncertainty. In a Bayesian context, DuMouchel [5] and Carlin [61] superimpose the posterior distribution and the likelihood for the among-study variance, respectively, on such a plot.

As a second example of a systematic sensitivity analysis, Copas and Shi illustrate the dependence of the summary effect estimate on different hypothetical models for publication bias [72]. They address a selection model in which the probability that a study is available depends on an overall publication probability and on the study's standard error, as well as on the estimated magnitude of effect. They illustrate how the overall meta-analysis estimate varies according to different values for the first two of these unobservable quantities.

As a third example, Elvik extracts confidence intervals for summary effects in a 'leave-one-out' sensitivity analysis to try and detect outlying studies, and plots these against the statistical information remaining in each meta-analysis [18]. The most influential (in terms of weight) will appear towards the lower end of the statistical information axis, and outlying studies will be identifiable by departures on the other axis, since the summary result on omission of an outlying study will differ from the summary results.

Copyright © 2010 John Wiley & Sons, Ltd.

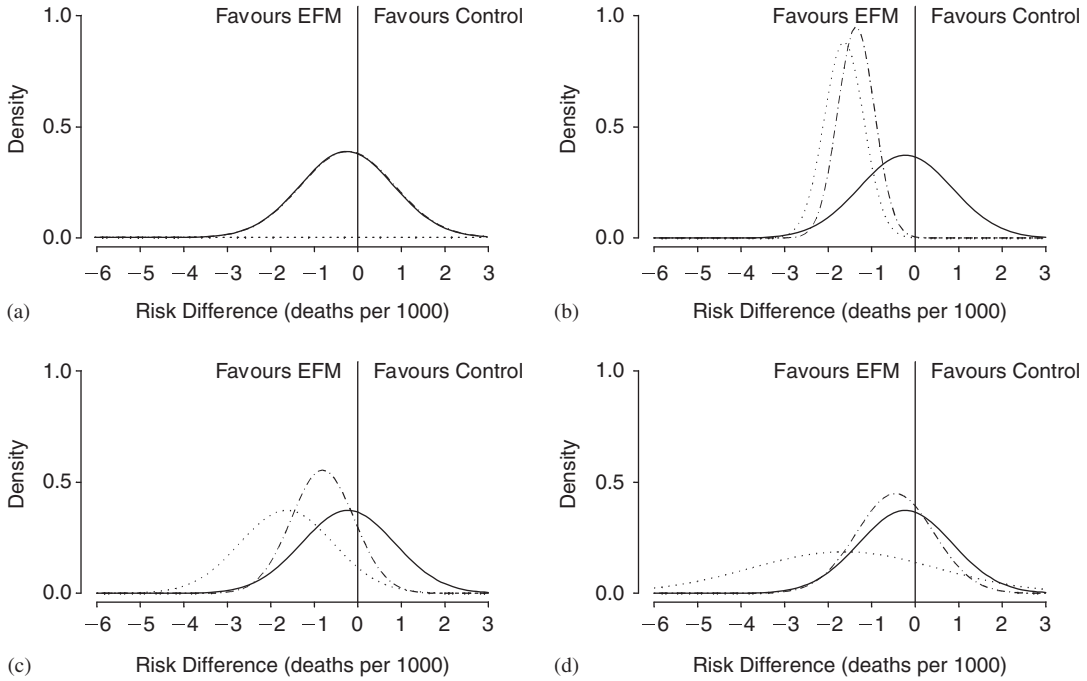

Figure 9. Likelihood and posterior distributions. Graphical representation of the prior beliefs about the effect of electronic fetal heart rate monitoring in perinatal death, using (a) vague prior, (b) naive prior, (c) equivalent prior, and (d) sceptical prior. The dotted line represents the prior, the full line the likelihood and the intermittent series of dots and lines represent the posterior [68]. Copyright Sage Publications. Reproduced with permission.

9. Plots of likelihoods and posterior distributions

9.1. Plotting likelihoods

Most meta-analyses are implemented simply as weighted averages, and the statistical theory that underlies them is often overlooked. The data for each study and the results of each meta-analysis can each be expressed as likelihoods, which reflect how likely various parameter values are (under a particular statistical model) given the observed data. In the standard meta-analysis method, in which an effect estimate from each study is assumed to be normally distributed, the likelihood from each study is a normal distribution. A plot of all of these likelihoods can look very busy but helps to determine the general tendency of the studies [49]. These may be integrated into a forest plot (Barrowman and Myers' 'raindrop plot') as described in Section 2.2. Likelihood plots for key parameters of interest are sometimes found in statistical papers on meta-analysis (for example, for the summary estimate [21, 49, 73], the heterogeneity variance [21, 73] and the joint likelihood of both [73]), but seldom in applied examples.

9.2. Bayesian meta-analyses

In Bayesian meta-analyses, likelihoods are combined with prior distributions to produce posterior distributions for each unknown quantity. We have already mentioned some displays of Bayesian meta-analysis results in Sections 2.2 and 8. Posterior distributions can also be plotted in a similar way to likelihoods in a classical analysis. An effective way to convey the impact of prior distributions is to plot them along with likelihoods and posterior distributions. We include an example from Sutton and Abrams in Figure 9 [68].

10. Discussion

Graphical presentation of results is an important part of a meta-analysis, as it is for almost any statistical analysis. We have described numerous options covering many facets of meta-analysis, and offered specific suggestions for the four that we consider to be the most generally useful, namely forest plots, funnel plots, Galbraith plots and L'Abbé plots. Our recommendation that the forest plot is a generally useful graph concurs with other recent commentaries [6, 7], and we propose that it continue to be the first choice when it is viable. We acknowledge that the use of funnel plots is more contentious, and they may be misleading for small numbers of studies [35]. Galbraith plots may be under-used given their mathematical attractions and ability to be informative about heterogeneity [6], but further work is warranted to assess their ease of interpretation in practice. L'Abbé plots (and related ROC plots) are appropriate only for studies comparing two groups, but for this they are very informative. Among

Copyright © 2010 John Wiley & Sons, Ltd.

the other plots, we consider normal probability plots and plots of systematic sensitivity analyses to be among the most useful, but as means of exploration rather than presentation of results.

Software options for most of these graphs are numerous, although restricted for the most part to applications or scripts created specifically for meta-analysis. Some examples that produce multiple types of plots include *Comprehensive Meta-analysis* (forest plots, funnel plots, bar charts of weights and residuals, scatter plots for meta-regression) [74], *MIX* (forest plots, funnel plots, L'Abbé plots, normal probability plots, Galbraith plots, Baujat plots and others) [29], *Review Manager* (forest plots, funnel plots, ROC plots) [75] and an extensive series of macros available for Stata [76].

Whenever graphs are used, basic principles should be followed [77], including the use of sensible axes, clear axis labels, adequate line thickness, adequate symbol sizes and unambiguous annotation and legends. We endorse a series of suggestions for meta-analyses made recently by Borman and Grigg, namely that good graphics principles can be applied also to tables; that graphs should emphasize the big picture, encourage the reader to make appropriate comparisons, and be selected on the basis of the purpose of the analysis and the nature of the data; and that the statistical precision of individual results should be reflected [7].

Acknowledgements

We are grateful to members of the Cochrane Statistical Methods Group for contributions to the document from which this paper evolved, particularly Doug Altman, Deborah Ashby, Jon Deeks, Gordon Dooley, Diana Elbourne, Sally Hollis, Steff Lewis, Keith O'Rourke, Jonathan Sterne and Simon Thompson. We thank an editor, an associate editor and two referees for helpful comments. This work was undertaken while Judith Anzures-Cabrera was working at the MRC Biostatistics Unit, and was funded by MRC grant number U.1052.00.011.

References

1. Barnard FR. *One Picture is Worth Ten Thousand Words*. Printers' Ink: New York, 1927; 114–115.
2. Borenstein M, Hedges LV, Higgins JPT, Rothstein H. *Introduction to Meta-analysis*. Wiley: Chichester, 2009.
3. Light RJ, Singer JD, Willett JB. The visual presentation and interpretation of meta-analyses. In *The Handbook of Research Synthesis*, Cooper H, Hedges LV (eds), Chapter 28. Russell Sage Foundation: New York, 1994; 439–445.
4. Ferrer RL. Graphical methods for detecting bias in meta-analysis. *Family Medicine* 1998; **30**:579–583.
5. DuMouchel W, Normand SL. Computer-modeling and graphical strategies for meta-analysis. *Meta-analysis in Medicine and Health Policy*, Stangl DK, Berry DA (eds). Marcel Dekker: New York, 2000; 127–178.
6. Bax L, Ikeda N, Fukui N, Yaju Y, Tsuruta H, Moons KG. More than numbers: the power of graphs in meta-analysis. *American Journal of Epidemiology* 2009; **169**:249–255.
7. Borman GD, Grigg JA. Visual and narrative interpretation. In *The Handbook of Research Synthesis and Meta-analysis* (2nd edn), Cooper H, Hedges LV, Valentine JC (eds), Chapter 26. Russell Sage Foundation: New York, 2009; 497–519.
8. Higgins J, Cochrane Statistical Methods Group. Considerations and recommendations for figures in Cochrane reviews: graphs of statistical data, 2003. Available from: http://www.cochrane.org/resources/handbook/ (accessed 1 January 2010).
9. DerSimonian R, Kacker R. Random-effects model for meta-analysis of clinical trials: an update. *Contemporary Clinical Trials* 2007; **28**:105–114.
10. Lewis S, Clarke M. Forest plots: trying to see the wood and the trees. *British Medical Journal* 2001; **322**:1479–1480.
11. Egger M, Davey Smith G, Phillips AN. Meta-analysis: principles and procedures. *British Medical Journal* 1997; **315**:1533–1537.
12. Moja L, Moschetti I, Liberati A. Understanding systematic reviews: the meta-analysis graph (also called 'forest plot'). *Internal and Emergency Medicine* 2007; **2**:140–142.
13. Ried K. Interpreting and understanding meta-analysis graphs. A practical guide. *Australian Family Physician* 2006; **35**:635–638.
14. Schriger DL, Altman DG, Vetter JA, Heafner T, Moher D. Forest plots in reports of systematic reviews: a cross-sectional study reviewing current practice. *International Journal of Epidemiology* 2010; DOI: 10.1093/ije/dyp370.
15. Galbraith RF. A note on graphical presentation of estimated odds ratios from several clinical trials. *Statistics in Medicine* 1988; **7**:889–894.
16. Sanders S, Glasziou PP, Del Mar CB, Rovers M. Antibiotics for acute otitis media in children. *Cochrane Database of Systematic Reviews* 2004; (1):CD000219.
17. DuMouchel W. Predictive cross-validation of Bayesian meta-analyses. *Bayesian Statistics 5*, Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds). Oxford University Press: Oxford, 1996; 107–127.
18. Elvik R. Evaluating the statistical conclusion validity of weighted mean results in meta-analysis by analysing funnel graph diagrams. *Accident Analysis and Prevention* 1998; **30**:255–266.
19. Ye Z, Liu EHC, Higgins JPT, Keavney Bd, Lowe GDO, Collins R, Danesh J. Seven haemostatic gene polymorphisms in coronary disease: meta-analysis of 66 155 cases and 91 307 controls. *The Lancet* 2006; **367**:651–658.
20. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research* 2001; **10**:277–303.
21. Van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Statistics in Medicine* 2002; **21**:589–624.
22. Smith TC, Spiegelhalter DJ, Thomas A. Bayesian approaches to random-effects meta-analysis: a comparative study. *Statistics in Medicine* 1995; **14**:2685–2699.
23. Higgins JPT, Thompson SG, Spiegelhalter DJ. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society Series A* 2009; **172**:137–159.
24. Lau J, Antman EM, Jimenez-Silva J, Kupelink B, Mosteller SF, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine* 1992; **327**:248–254.

Copyright © 2010 John Wiley & Sons, Ltd.

25. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *Journal of the American Medical Association* 1992; **268**:240–248.
26. Lau J, Schmid CH, Chalmers TC. Cumulative meta-analysis of clinical trials: builds evidence for exemplary medical care. *Journal of Clinical Epidemiology* 1995; **48**:45–57.
27. Pogue J, Yusuf S. Overcoming the limitations of current meta-anlysis of randomised controlled trials. *The Lancet* 1998; **351**:47–52.
28. Ioannidis JPA, Contopoulos-Ioannidis DG, Lau J. Recursive cumulative meta-analysis: a diagnostic for the evolution of total randomized evidence from group and individual patient data. *Journal of Clinical Epidemiology* 1999; **52**:281–291.
29. Bax L, Yu LM, Ikeda N, Tsuruta H, Moons KG. Development and validation of MIX: comprehensive free software for meta-analysis of causal research data. *BMC Medical Research Methodology* 2006; **6**:50.
30. Barrowman NJ, Myers RA. Raindrop plots: a new way to display collections of likelihoods and distributions. *American Statistician* 2003; **57**:268–274.
31. Light RJ, Pillemer DB. *Summing Up*: *The Science of Reviewing Research*. Harvard University Press: Cambridge, MA, 1984.
32. Sterne JAC, Egger M, Davey Smith G. Investigating and dealing with publication bias and other biases. *Systematic Reviews in Health Care*: *Meta-analysis in Context*, Egger M, Davey Smith G, Altman DG (eds). BMJ Publication Group: London, 2001; 189–208.
33. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 1997; **315**:629–634.
34. Renehan AG, Tyson M, Egger M, Heller RF, Zwahlen M. Body-mass index and incidence of cancer: a systematic review and meta-analysis of prospective observational studies. *The Lancet* 2008; **371**:569–578.
35. Lau J, Ioannidis JPA, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *British Medical Journal* 2006; **333**:597–600.
36. Tang JL, Liu JL. Misleading funnel plot for detection of bias in meta-analysis. *Journal of Clinical Epidemiology* 2000; **53**:477–484.
37. Munafo MR, Clark TG, Flint J. Assessing publication bias in genetic association studies: evidence from a recent meta-analysis. *Psychiatry Research* 2004; **129**:39–44.
38. Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *Journal of Clinical Epidemiology* 2001; **54**:1046–1055.
39. Terrin N, Schmid CH, Lau J. In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology* 2005; **58**:894–901.
40. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *Journal of Clinical Epidemiology* 2008; **61**:991–996.
41. Sterne JAC, Egger M, Moher D. Addressing reporting biases. In *Cochrane Handbook for Systematic Reviews of Interventions* Higgins JPT, Green S (eds), Chapter 10. Wiley: Chichester, U.K., 2008; 297–333.
42. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 2001; **56**:455–463.
43. Higgins JPT, Altman DG (eds). Assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions*. Higgins JPT, Green S (eds), Chapter 8. Wiley: Chichester, U.K., 2008; 187–241.
44. Galbraith RF. Some applications of radial plots. *Journal of the American Statistical Association* 1994; **89**:1232–1242.
45. Whitehead A, Whitehead J. A general parametric approach to the meta-analysis of randomised clinical trials. *Statistics in Medicine* 1991; **10**:1665–1677.
46. Thompson SG. Controversies in meta-analysis: the case of the trials of serum cholesterol reduction. *Statistical Methods in Medical Research* 1993; **2**:173–192.
47. Copas J, Lozada-Can C. The radial plot in meta-analysis: approximations and applications. *Journal of the Royal Statistical Society Series C* (*Applied Statistics*) 2009; **58**:329–344.
48. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Annals of Internal Medicine* 1987; **107**:224–233.
49. Van Houwelingen HC, Zwinderman KH, Stijnen T. A bivariate approach to meta-analysis. *Statistics in Medicine* 1993; **12**:2273–2284.
50. Marret E, Remy C, Bonnet F, Postoperative Pain Forum Group. Meta-analysis of epidural analgesia versus parenteral opioid analgesia after colorectal surgery. *British Journal of Surgery* 2007; **94**:665–673.
51. Jimenez FJ, Guallar E, Martinmoreno JM. A graphical display useful for meta-analysis. *European Journal of Public Health* 1997; **7**:101–105.
52. Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Statistics in Medicine* 2002; **21**:1575–1600.
53. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *Journal of Clinical Epidemiology* 1995; **48**:119–130.
54. Langlotz CP, Sonnad SS. Meta-analysis of diagnostic procedures: a brief overview. *Academic Radiology* 1998; **5**:S269–S273.
55. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002; **21**:1539–1558.
56. Walker AM, Martin-Moreno JM, Artalejo FR. Odd man out: a graphical approach to meta-analysis. *American Journal of Public Health* 1988; **78**:961–966.
57. Pladevall-Vila M, Delclos GL, Varas C, Guyer H, Brugués-Tarradellas J, Anglada-Arisa A. Controversy of oral contraceptives and risk of rheumatoid arthritis: meta-analysis of conflicting studies and review of conflicting meta-analyses with special emphasis on analysis of heterogeneity. *American Journal of Epidemiology* 1996; **144**:1–14.
58. Rifat SL. Graphic representations of effect estimates: an example from a meta-analytic review. *Journal of Clinical Epidemiology* 1990; **43**:1267–1269.
59. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine* 1998; **17**:841–856.
60. Wang MC, Bushman BJ. Using the normal quantile plot to explore meta-analytic data sets. *Psychological Methods* 1998; **3**:46–54.
61. Carlin JB. Meta-analysis for 2 x 2 tables: a Bayesian approach. *Statistics in Medicine* 1992; **11**:141–158.
62. Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiologic Reviews* 1987; **9**:1–30.
63. Baujat B, Mahe C, Pignon JP, Hill C. A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. *Statistics in Medicine* 2002; **21**:2641–2652.
64. Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine* 2002; **21**:1559–1574.
65. Thompson SG. Why and how sources of heterogeneity should be investigated. In *Systematic Reviews in Health Care*: *Meta-analysis in Context*, Egger M, Davey Smith G, Altman DG (eds), Chapter 9. BMJ Books: London, 2001; 157–175.
66. Bailar III JC. The practice of meta-analysis. *Journal of Clinical Epidemiology* 1995; **48**:149–157.

Copyright © 2010 John Wiley & Sons, Ltd.

67. Berkey CS, Hoaglin DC, Mosteller F, Colditz GA. A random-effects regression model for meta-analysis. *Statistics in Medicine* 1995; **14**:395–411.
68. Sutton AJ, Abrams KR. Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research* 2001; **10**:277–303.
69. Lau J, Ioannidis JPA, Schmid CH. Summing up evidence: one answer is not always enough. *The Lancet* 1998; **351**:123–127.
70. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. *The Lancet* 2002; **360**:1903–1913.
71. Olkin I. Diagnostic statistical procedures in medical meta-analyses. *Statistics in Medicine* 1999; **18**:2331–2341.
72. Copas J, Shi JQ. Meta-analysis, funnel plots and sensitivity analysis. *Biostatistics* 2000; **1**:247–262.
73. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. *Statistics in Medicine* 1996; **15**:619–629.
74. Borenstein M, Hedges L, Higgins J, Rothstein H. *Comprehensive Meta-analysis Version 2 [Computer program]*. Biostat: Englewood, NJ, 2005.
75. *Review Manager* (*RevMan*) [*Computer program*] *Version 5.0*. Nordic Cochrane Centre: Copenhagen,
76. Sterne JAC (ed.). *Meta-Analysis in Stata*: *An Updated Collection from the Stata Journal*. Stata Press: College Station, TX, 2009.
77. Cleveland WS. *The Elements of Graphing Data*. Hobart Press: Summit, NJ, 1994.

Copyright © 2010 John Wiley & Sons, Ltd.