Eur J Vasc Endovasc Surg (2010) 40, 669–677

REVIEW

Quality Control in Systematic Reviews and Meta-analyses

M.J. Bown [a,*], A.J. Sutton [b]

[a] *The University of Leicester, Department of Cardiovascular Sciences, Robert Kilpatrick Clinical Sciences Building, Leicester Royal Infirmary, Leicester LE2 7LX, United Kingdom*
[b] *The University of Leicester, Department of Health Sciences, United Kingdom*

Submitted 7 April 2010; accepted 16 July 2010
Available online 21 August 2010

KEYWORDS
Systematic review;
Meta-analysis;
Guidelines

Abstract Systematic reviews and meta-analyses are being submitted to, and being published by biomedical journals with increasing frequency. In order to maintain the utility of such publications and avoid misguidance it is important that these studies are conducted to a high standard. This article aims to provide guidance both for those researchers undertaking and reporting such studies and for the readers of such articles. Details of a suggested method for conducting a systematic review are given, including methods for literature searches, data abstraction and data extraction followed by a brief overview of common methods used for meta-analyses and the interpretation of the results of meta-analysis.

Introduction

Evidence synthesis via the publication of systematic reviews, particularly when these contain meta-analytical combination of data, is now an established part of the repertoire of many scientific journals. The number of systematic reviews and meta-analyses has grown steadily over the last two decades (Fig. 1).[1] Systematic reviews have significant benefits over conventional reviews in that all available data is presented, not just that chosen by the authors. However to ensure that this is the case, authors of systematic reviews must be sure to perform their review with rigorous attention to detail and report their methods to enable adequate scrutiny of their conclusions. In the reporting of the results from this type of research there is the potential for erroneous conclusions to be presented by the author or, drawn by the reader. If incorrect conclusions are used to guide clinical decision making this may lead to ineffectual or harmful treatment being administered. It should also be remembered that the results from any systematic review or meta-analysis are only as good as the source data upon which they are based. Without proper caution when combining studies potentially erroneous conclusions can be reached. This is a particular problem if the quality of the contributing studies is not considered or analysed.

* Corresponding author. Tel.: +44 1162523252; fax: +44 1162523179.
E-mail address: m.bown@le.ac.uk (M.J. Bown).

1078-5884/$36
doi:10.1016/j.ejvs.2010.07.011

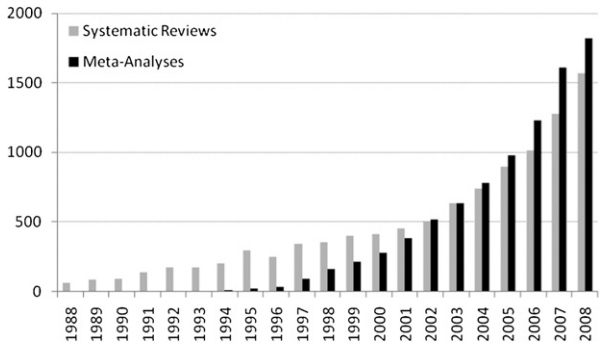

Figure 1 The numbers of publications each year from 1988 to 2008 with the words "systematic review" (grey bars) or "meta-analysis" (black bars) in the title. Data retrieved from ISI Web of Knowledge based on a search of MEDLINE® (1950-present).

There is often confusion regarding the precise definition of what is a systematic review and what is a meta-analysis. Both use bibliographic methods to obtain as complete as possible a dataset of individual studies that have examined the same hypothesis or related hypotheses and then combine the data in order to reach an overall interpretation of the combined data. However, a meta-analysis involves the mathematical combination of the results from the source data whilst a systematic review does not. The processes of mathematical combination used for meta-analyses have their own particular set of methods and standards for reporting. It is the reporting of these methods and results that usually makes up a large proportion of the difference between a systematic review and a meta-analysis.

The potential for misinterpretation of data is particularly problematic in meta-analyses. Because of the nature of the statistical techniques used in meta-analyses the majority of readers will not be familiar with interpreting the significance of the presented data and in particular, the interpretation of the various quality control measures (such as publication bias assessment and the results of sensitivity analyses). Furthermore, there is evidence to support caution in the assumption that all meta-analyses provide better evidence than individual trials. In 2005 a meta-analysis of peri-operative beta blockade suggested that there was a mortality benefit for this treatment (albeit with increased side-effect profile)[2] and this evidence led to recommendations for the prescription of peri-operative beta-blockers in patients undergoing non-cardiac surgery[3] (although only approximately 40% of the patients in this study underwent vascular surgical procedures). However, since the publication of this meta-analysis, the result of a large randomised controlled trial with twice as many participants as the combined meta-analysis has become available[4] and this suggests that peri-operative beta blockade is actually harmful — a finding contradictory to the result of the previous meta-analysis. Readers should also be aware that there is a hierarchy of evidence in meta-analyses. The outputs from two meta-analyses may appear similar but their quality is based upon the contributing studies. A meta-analysis of several randomised controlled trials with identical methods is of far greater quality than one combining many observational studies with variable inclusion/exclusion criteria, time periods or treatment types.

The aim of this article is twofold. Firstly, we wish to provide guidance to those researchers who are considering undertaking any form of evidence synthesis whether this is a systematic review, meta-analysis of comparative studies or a meta-analysis of observational studies. Secondly, this article aims to de-complicate the interpretation of meta-analytical results and the methods used to check the strength of the associations seen, and to enable the occasional reader to critically appraise the evidence presented in systematic reviews for their own benefit. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) statement provides specific guidance on the conduct of systematic reviews and meta-analyses (with a focus on meta-analyses of randomised controlled trials) and the MOOSE (Meta-analysis Of Observational Studies in Epidemiology) statement is specifically designed for meta-analyses of observational studies.[5,6] The Cochrane Handbook for Systematic Reviews of Interventions[7] is an essential resource for those conducting systematic reviews and meta-analyses, as is the United Kingdom National Institute for Health Research Centre for Reviews and Dissemination.[8] It is not the authors' intention to reproduce the work performed by these groups and the available guidelines contained within these works should be used in conjunction with this article by authors preparing their own systematic reviews and meta-analyses. Whilst diagnostic research can be assessed through meta-analysis this area is outside the scope of this review, full details of how to perform this type of study can be found elsewhere.[9]

Part 1: Performing and Reporting a Systematic Review

Planning a systematic review

The choice of subject for any review should be based on either clinical or scientific need for the review. The review should answer specific questions that need to be carefully defined at the outset of the exercise and clearly stated in the opening stages of any report. A useful method for defining the exact clinical question(s) to be answered is to categorise and specify the patient group, the intervention, the comparison and the measured outcome. Simply trawling the literature and 'seeing what comes out' is more likely to add un-necessary outcomes of little relevance rather than clarify available evidence. The overall value of any review can only be related to the source information. Ideally, there should be a good body of evidence relating to the subject, if there is not then a review is unlikely to add significantly to the evidence in that area. This can only realistically be determined by performing a literature search. One of the reasons for the recent increase in the number of systematic reviews published (Fig. 1) may be that multiple reviews of the same subject are often now performed, as can be seen for carotid stenting.[10–13] Whilst the updating of a previously published systematic review is to be encouraged,[12] this should only be performed if a large enough body of new

evidence has emerged since the previous publication and redundancy should be avoided.

Literature searches and study selection

Good quality systematic reviews necessitate good quality literature searches, and accurate reporting of these searches. Searching of single databases will only identify a maximum of one third of all relevant articles and searching multiple databases still only identifies half of all available articles.[14] The main databases used are Medline, Embase, and The Cochrane Library. Common portals/servers used to access these databases are PubMed and Ovid although many others exist. It is important to appreciate that one portal may only search a subset of any particular database and for the most comprehensive searches multiple portals in addition to multiple databases should be used. The services of a clinical librarian or information specialist should be employed to enhance the quality of the literature search.

The use of multiple searches using multiple search terms, different combinations of search terms and search term synonyms also improves the effectiveness of an electronic literature search. Where a search term is spelled differently in different countries these alternative spellings of search terms should be included as separate searches. Study selection and exclusion criteria will affect the literature base for a systematic review. These both need to be clearly stated in any report. The criteria used for inclusion will often determine whether a meta-analysis can be performed or not. If inclusion criteria are too strictly defined then there will not be enough data to allow meaningful combination of results. Conversely, if inclusion criteria are non-specific a large dataset of non-homogenous studies will be obtained. Since meta-analyses perform better in homogenous datasets this strategy may result in meaningless output. However, this second strategy does allow for the selection of subsets of studies that do demonstrate homogeneity and will ensure that all relevant/possible information is obtained from the review.

Many authors will exclude articles from a systematic review or meta-analysis based upon language of publication on the basis of practicality (e.g. translational expenses) and since the majority of biomedical literature is published in English language journals this is thought to include the majority of relevant articles. However this approach is possibly misconceived. If Medline (1950 to date) is searched for the keyword "vascular surgery" 6088 articles are identified. When non-English language articles are excluded only 4228 articles remain (search performed 26th September 2009 using OvidSP_UI02.02.00.156). There is evidence that this practice may have significant effects on the results of a systematic review and currently should not be accepted as a routine practice, particularly when meta-analysis of clinical trials is being performed.[15] In one study a meta-analysis restricted to English language literature that resulted in a treatment non-recommendation was re-analysed including the non-English language literature. This significantly altered the results of the meta-analysis and demonstrated a positive treatment effect.[16] However, some authors suggest that language bias in most meta-analyses may not be as important as first thought - although there is an effect in a minority of studies.[17] Because of the uncertainty surrounding the true effect of language bias the exclusion of studies based on language should be fully justified, the excluded studies should be clearly cited and, in the case of meta-analyses of randomised trials, this practice should not be encouraged. Publication language should be considered during the sensitivity analysis phase of a meta-analysis. Other exclusion criteria need to similarly be stated and carefully considered prior to their adoption. In cases where authors are uncertain regarding the justification of any potential exclusion criteria these should be not be used but instead applied to the study at the time of performing a sensitivity analysis (vide infra).

Careful analysis of the studies obtained from a literature search has to be made to look for possible duplicate publication of the same data. Often this is not due to overt publication of the same data in different journals but, more commonly, due to multiple publications over time from a single centre with later publications including the data from earlier publications. To screen for this it is often useful to extract the location and institution where studies were performed and the start and finish dates of the study.

Systematic reviews are performed at a single specific point in time. By their nature they become outdated as soon as new evidence in their subject area becomes available, which is often prior to the publication of the review itself. The date at which the literature search was performed needs to be reported precisely. This then enables future researchers to update the systematic review by simply repeating the search from that time point. This also underlines the need for accurate reporting of the overall search strategy. This ability to update systematic reviews is central to the principles that underlie the Cochrane collaboration.[18,19] However it should also be noted that there is a danger of updating systematic reviews until a positive result is identified[20] and therefore updating should continue beyond this point (there is the chance that with the addition of future data a result may become non-significant).

Good examples of literature search reporting can be found in the systematic reviews by Rerkasem et al., Culwell et al. and Rebollo Aguirre et al.[21–23]

Data abstraction and extraction

Data abstraction (the selection of data items to be extracted from each contributing article) and data extraction form an essential part of any systematic review. Often the data abstraction will depend upon the aims of the systematic review and often form part of the selection criteria for study inclusion. For example, if a systematic review sets out to examine the effect of one particular therapeutic intervention upon a clinical outcome the data abstraction will include the number of patients studied, the number receiving the intervention and the clinical outcomes in each group. There would be little point in selecting articles for inclusion in the review if this data was not available and these variables become part of the essential dataset. If a meta-analysis is to be included in the review the primary data must include a measure of the

dispersion of any outcome estimate to enable statistical combination. The data necessary to conduct a meta-analysis varies according to the outcome measure being compared. Examples for several common scenarios are shown in Table 1.

The gold standard for data extraction is for multiple individuals to independently perform the data extraction, compare results and resolve any discrepancies by consensus. This approach necessitates good data abstraction prior to extraction since repeating data extraction in this scenario is especially time-consuming. It is acceptable to perform double data extraction on a proportion of the data, check the level of agreement and if this is high, continuing from this point with single individual data extraction. Alternative, less robust, strategies include single researchers conducting data extraction at two separate time points or (the weakest strategy) single researchers conducting data extraction on a single occasion.

Part 2: Process and Interpretation of Meta-analyses

The first question it is necessary to answer when considering the incorporation of a meta-analysis into a systematic review is whether there is enough individual study data to warrant statistical combination. There is no fixed number of studies or combined number of individuals that can be used as a threshold to aid this decision. A small number of studies with similar outcomes will provide a tidy statistical result but one has to be sure that the amount of evidence being combined is adequate to address the thesis posed. The combination of larger numbers of studies (more than 20) permits more detailed analysis and more rigorous quality control of the results obtained through sensitivity analyses and publication bias assessments.

An important concept in all systematic reviews, including meta-analyses is that of heterogeneity. This term is often used in different circumstances at different times during such studies. Heterogeneity refers to the amount of variability seen between studies. Clinical heterogeneity can arise from differences in the patient populations that were examined in different studies or differences in treatments used. Methodological heterogeneity is a result of differences in experimental design between studies. Statistical heterogeneity is most frequently quoted since a measure for this is derived as part of a meta-analysis and often used to guide or modify the statistical analyses performed. Whilst statistical heterogeneity is most commonly referred to, it is probably more important to consider clinical and methodological heterogeneity prior to performing any meta-analytical statistical combination of results. The reason behind this is that if it is considered that there is a high level of clinical and/or methodological heterogeneity the decision to proceed with a meta-analysis has to be reviewed.

Data types and outcome measures

Meta-analyses may be used to combine many different types of outcome data. Dichotomous, categorical,

Table 1 Data to be extracted from individual studies to permit meta-analysis.

Type of meta-analysis		Data needed for meta-analysis:	
Comparison of treatments	Categorical outcome	Number of outcome events and number of observations in each treatment group.	Usually provided, can only be derived if all individual data listed in contributing study
	Continuous outcome	Mean, standard deviation, and number of observations for each treatment, OR, the mean difference between outcomes and standard error for this difference.	
Combination of observational studies	Categorical outcome	Number of outcome events and number of observations.	
	Continuous outcome	Mean, standard deviation, and number of observations for each treatment	

continuous, survival, event rate or a combination of data types may be combined (pooled) using a variety of different statistical methods. Many outcome measures are expressed as ratios (odds, risk, hazard). These measures are centred on a value of 1 which implies that there is no difference between the treatment assessed and no intervention or, in the case of observational studies, an even likelihood of the event occurring by chance alone. It is important to note that there is a difference between Odds Ratio (OR) and Relative Risk (RR). Both are acceptable measures of effect size. Odds Ratios are more useful for descriptive purposes whilst for interventional treatment comparisons Relative Risk is favoured.[24]

Meta-analytical models

Meta-analyses are usually conducted using either 'fixed-effects' or 'random-effects' models (and this should be stated by the authors). These models make different assumptions about the overall outcome being assessed. For fixed-effects meta-analyses it is assumed that there is a fixed value for the outcome being measured in the whole population and that each study contributing to the meta-analysis is estimating this fixed value in its own sample drawn from the total population. For random-effects meta-analyses it is assumed that the outcome being estimated is not of a fixed value in every population, but the values are related between populations. Each study is therefore estimating the value in its own study population and these populations are randomly sampled from the overall population. The use of a fixed-effects model is indicated where there is low heterogeneity (variability) in the individual studies contributing to the meta-analysis, whilst random-effects models are used when the contributing studies have heterogeneous estimates of the outcome measure being combined.

The decision to use a random-effects or fixed-effects model is usually based on the degree of heterogeneity between the individual studies. If the estimates of the outcome of interest are approximately the same across all studies the between-study-heterogeneity is said to be low and it is therefore appropriate to use a fixed-effect model. If the individual studies show a wide variation in the outcome estimates (heterogeneity) then a random-effects model is appropriate. Heterogeneity is usually assessed by using a Cochran's chi-squared test and *P*-values of less than 0.1 (and not the usual 0.05) indicate significant heterogeneity. This approach is limited however when the number of studies is large as heterogeneity assessed by a simple chi-squared test is often seen. This is largely because of the high power to be able to detect even small degree of heterogeneity with a large sample size (number of studies). In this situation it may be that a chi-squared test may be detecting heterogeneity that is not clinically meaningful and this should be considered before dismissing a result due to a high degree of heterogeneity. An alternative approach is to use an H, R, tau^2(τ^2) or I^2 value.[25] Unfortunately there remains significant disagreement amongst the statistical community regarding the interpretation and use of these statistics.[26–28] For the non-expert a Cochran's Q *P*-value of <0.1 or an I2 value greater than 50% should be considered to represent significant between-study-heterogeneity and a random-effects model should be used. It should also be noted that if an author considers there to be clinical heterogeneity between studies a random-effects model should be used irrespective of the results of any statistical test for heterogeneity.

Reporting of results

The reporting of results from meta-analyses should enable the reader to examine the individual study data without recourse to obtaining the references themselves. This should include the numbers of patients, the types and numbers of interventions and the outcomes in each experimental/observed group. Detail should also be given of those criteria that were used to determine sub-groups, selection for sensitivity analyses or meta-regression analyses.

Graphical presentation of the results from meta-analyses is routinely presented as forest plots (Fig. 2). These are a combined tabular and graphical representation of each individual study contributing to the meta-analysis. The study identifier and outcome estimate for that study together with a measure of outcome uncertainty (usually a confidence interval) and the weight of that study in the analysis are usually presented on the plot and a graphical

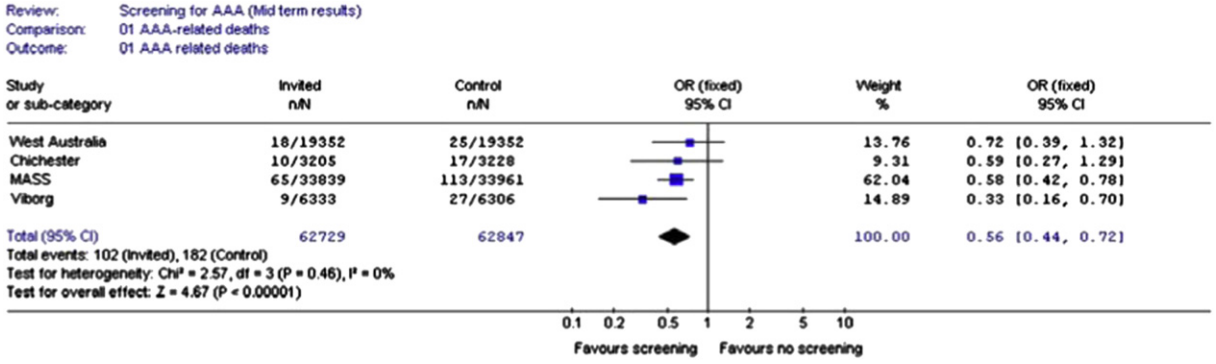

Figure 2 A forest plot of outcome data from a meta-analysis of 4 randomised trials of screening for Abdominal Aortic Aneurysm in men aged 64 to 83. Note the logarithmic scale and the vertical line at the threshold between outcome favouring treatment (in this case, screening) or control therapies. Reproduced with permission from: Lindholt JS. Norman P. Screening for abdominal aortic aneurysm reduces overall mortality in men. A meta-analysis of the mid- and long-term effects of screening for abdominal aortic aneurysms. European Journal of Vascular & Endovascular Surgery. 36(2):167—71, 2008.

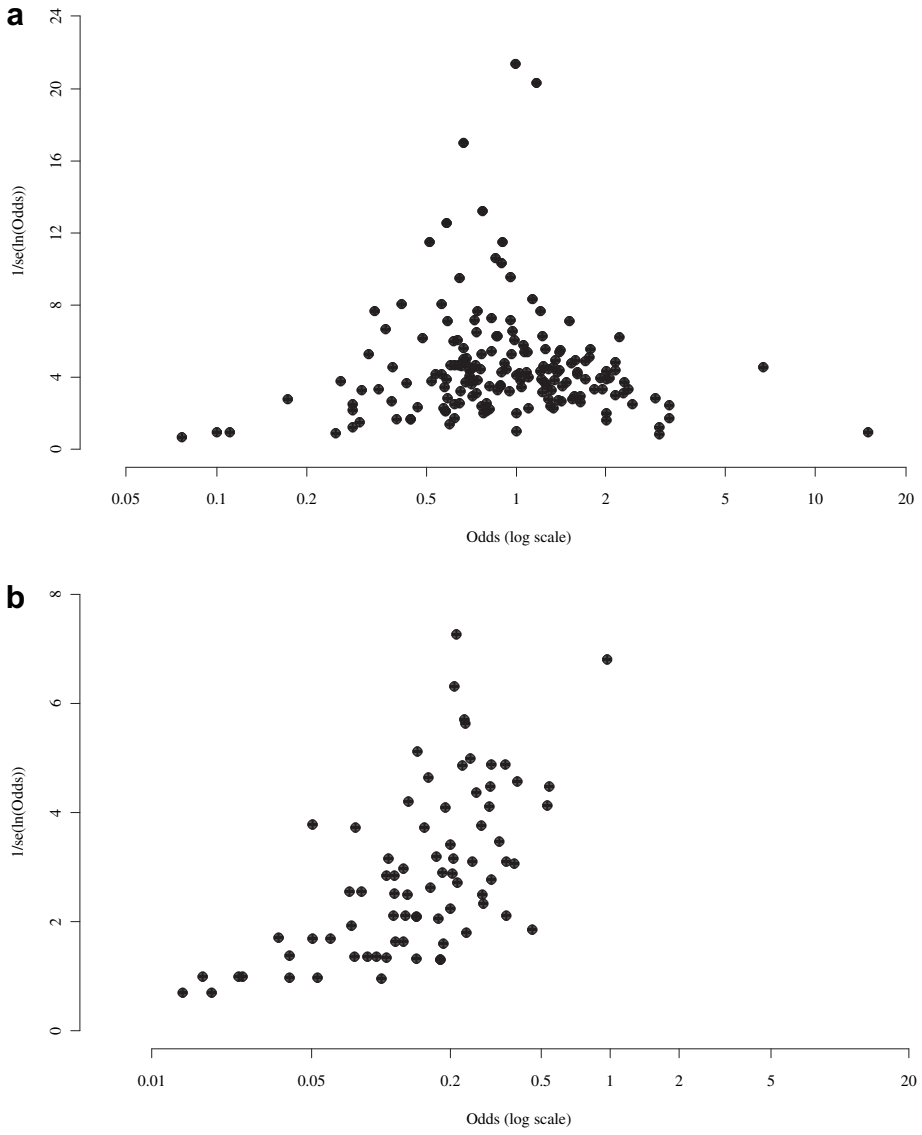

Figure 3 a) A funnel plot demonstrating reasonable symmetry from a meta-analysis of 171 studies. b) A funnel plot demonstrating marked asymmetry (77 studies). In this case the outcome measure was intra-operative mortality and the missing studies that should occupy the lower right part of the funnel represent those with high intra-operative mortality rates. Both figures reproduced with permission from Bown MJ, Sutton AJ, Bell PR, Sayers RD. A meta-analysis of 50 years of ruptured abdominal aortic aneurysm repair. Br J Surg. 2002 Jun;89(6):714—30.

representation with point size indicating the analysis weighting and error bars indicating the confidence interval. At the bottom of the plot the combined outcome estimate and model statistics are given on the left and a diamond on the graphical portion indicates the estimate, with its width representing the confidence interval of the estimate. A solid vertical line on the plot at an outcome measure equivalent to no difference between treatments or no difference from the outcome expected by chance is usually present. Occasionally an interrupted vertical line is marked from the centre of the combined outcome estimate. The distribution of the studies and their confidence intervals around the combined estimate can also provide visual clues as to the overall degree of heterogeneity in the analysis and which studies are responsible for this heterogeneity. Forest plots are often modified to include data from several distinct meta-analyses or to present data from sub-groups within meta-analyses.

Publication bias

The predisposition of journals to favour publication of positive reports over negative investigative findings[29] and the reticence of authors to publish poor outcomes (particularly in surgical specialities) leads to a high chance that a meta-analysis may be affected by these publication biases. It is therefore important to check for the presence of publication bias in any dataset used to perform a meta-analysis. Moreover, the assessment of potential publication bias is deemed to be an essential part of performing

meta-analyses[6] but despite this, it is often missing from reports.[30]

The most commonly used method to assess a dataset for potential publication bias is through the construction of funnel plots.[31] Funnel plots are scatter plots with one point for each study included in the meta-analysis. The outcome measure for each study is plotted on the *x*-axis against a measure of study accuracy on the *y*-axis (usually the same value used to weight that particular study in the meta-analysis although study size can be used). The highest weighted/largest/most accurate studies will align to the overall outcome seen in the meta-analysis. With decreasing size the individual study outcomes will deviate more from the overall meta-analytical outcome. If there is no publication bias present these smaller studies will be distributed evenly each side of the overall meta-analytical outcome and the plot will resemble an inverted funnel (Fig. 3a). If the funnel is asymmetric (Fig. 3b) this implies that there are studies 'missing' from the literature. This analysis can be assisted through the use of contour-enhanced funnel plots or statistical testing for asymmetry.[31,32] Depending on the position of the missing studies inferences can be drawn as to whether they represent studies with poor outcomes (as in Fig. 3b) or suppression of positive results. The precise causes of publication bias cannot be drawn from funnel plots but can include true publication bias (journals only publishing positive results), poor design in smaller studies, fraud and, in surgery, technical zealots producing good personal results that do not translate into the general population. It is important to remember that publication bias is not the only cause of funnel plot asymmetry.[33,34] Care should be taken when examining funnel plots constructed from small numbers of studies. For example in the study by Agarwal et al.[35] the funnel plot shows asymmetry in just 5 studies plotted but if the smallest study is removed the plot resumes a more symmetrical shape. Some authors construct funnel plots with the axes reversed[36] however this does not alter the method for interpretation. If there is evidence of publication bias it is possible to attempt corrections of the meta-analysis to compensate for these missing studies. Methods that have been used are 'trim and fill',[37] regression based methods,[38] and the exclusion the studies judged to be contributing to the asymmetry by visual assessment.[39]

Sensitivity and sub-group analyses

A sensitivity analysis is an important part of a meta-analysis as it aims to determine the robustness of the observed outcomes to the assumptions made in performing the analysis. Unfortunately this essential part of meta-analyses is often either not performed or reported. There is no set strategy for performing a sensitivity analysis. The underlying principle is to repeat the primary analysis with an altered dataset or statistical method to determine whether these changes have any effect on the combined outcome estimate. When altering the dataset, the choice of studies to add or remove is often based on assumptions of quality or study size and is usually at the author's discretion. It may be that a meta-analysis contains a mixture of randomised and non-randomised clinical trials or multi-centre and single-centre studies. Studies perceived to be of lower quality are removed and the analysis is then repeated. If

Table 2 Essential criteria to be included in systematic reviews and meta-analyses.

Systematic reviews	Meta-analyses[a]
Definition of intervention/technique/subject of systematic review	Outcomes listed
Literature search	Methods
Search terms and combinations	Software package
Databases used	Weighting method
Date of search	Heterogeneity assessment
Hand searching of references used?	Choice of fixed or random-effects model
Authors performing literature search	
(must be at least 2)	Characteristics of studies contributing to meta-analysis:
	Study-level inclusion/exclusion criteria
Inclusion and exclusion criteria explicitly defined	Individual study methodology
	Study participants
Data abstraction	Timescale
	Setting.
Authors performing data extraction	
	Results
Study flow diagram	Statement of main outcome
	Forest plot(s)
	Sensitivity analysis
	Publication bias assessment

All submissions must include a summary stating in 50 words or less what that study has added to the current knowledge of the subject addressed.

[a] Additional requirements in addition to those required for systematic reviews.

the analysis is robust then there should be little change in the overall outcome estimate. An alternative to removing studies based on quality criteria is to include studies excluded a priori during the data-abstraction phase of the meta-analysis. Sensitivity analyses can be performed by altering the statistical model (for example applying a fixed-effects model in the sensitivity analysis where a random-effects model has been used initially and vice versa).

Meta-analyses combine data by definition and in this process subtle effects or differences between individual studies may be lost in an attempt to aggregate enough data to permit analysis. Sub-group analyses attempt to pick out this lost information and are similar to sensitivity analyses in that the dataset is re-analysed except for the fact that rather than adding/removing studies the entire dataset is analysed but divided by some factor of interest to the authors. The danger with sub-group analyses is that the amount of primary data being put forward into each analysis is usually small and conclusions drawn from these smaller datasets have to be guarded.

There are many examples in the literature of meta-analyses where sensitivity analysis has not been performed. Some examples where sensitivity analyses have been performed can be found[35,40] however the adoption of this practice is by no means universal.

Meta-regression

In meta-analyses, as with many statistical analyses we wish to determine how an independent factor may affect our outcome measure. In conventional statistical techniques regression is used to determine the effect of one factor upon an outcome variable and a similar technique called meta-regression can be employed as part of a meta-analysis. Meta-regression is similar to sub-group analysis except the factors chosen are usually continuous, such as the publication date of each study contributing to the meta-analysis or the median age of patients in each study contributing to the meta-analysis. It should be noted that sub-group analysis could be performed using meta-regression techniques. If meta-regression is used for sub-group analysis formal statistical tests to determine differences between sub-groups are available. Meta-regression is particularly useful to determine whether an outcome changes with respect to the factor, for example time. It does suffer from the same limitations as sub-group analysis and requires a large volume of individual studies to make meaningful interpretations from the data. It should also be remembered that *P*-values quoted for the regression equation are based on the null hypothesis that the slope of the regression line is zero and significant outcomes simply suggest that there is a relationship between the factor and outcome being compared. Meta-regression usually assumes a linear relationship between the explanatory variable and the outcome measure.

Computer software

The conduct of meta-analysis is greatly facilitated by the use of statistical packages with meta-analysis functionality or stand-alone specialised applications.[41] Freely available user macros for the commercial package Stata[42] and the freely available package *R*[43] offer extensive functionality but require some programming knowledge of the underlying software platform. One of the most commonly used stand-alone packages is the software developed by the Cochrane Collaboration, Revman (freely available for academic use), which has an intuitive graphical user interface.[18] Other non-freeware packages include Comprehensive meta-analysis,[44] MIX 2.0[41,45] (although a less powerful freeware version is available) and StatsDirect[46] (the latter two provide front-ends to EXCEL).

Conclusions and guidelines for authors

We have attempted to clarify the processes behind evidence synthesis projects with particular reference to literature searching for systematic reviews and the interpretation of meta-analyses. The primary aim of this piece of work is to aid readers in the critical review of such articles but in doing so several quality criteria can be drawn for which authors submitting such articles to biomedical journals such as The European Journal of Vascular and Endovascular Surgery should consider. These are listed in Table 2. As stated above, these criteria should be used in conjunction with the guidelines already published by the international consensus committees[5,6] and are not intended to replace them.

Conflict of Interest

None.

Funding

None.

References

1 *Reuters ISI Web of Knowledge. [Internet]*. Reuters [accessed 26.09.2009]. Available: http://wokinfo.com/; 2009.
2 Devereaux PJ, Beattie WS, Choi PT, Badner NH, Guyatt GH, Villar JC, et al. How strong is the evidence for the use of perioperative beta blockers in non-cardiac surgery? Systematic review and meta-analysis of randomised controlled trials. *BMJ* 2005;**331**(7512):313—21.
3 Fleisher LA, Beckman JA, Brown KA, Calkins H, Chaikof EL, Fleischmann KE, et al. ACC/AHA 2007 guidelines on perioperative cardiovascular evaluation and care for noncardiac surgery: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 2002 Guidelines on Perioperative Cardiovascular Evaluation for Noncardiac Surgery) Developed in Collaboration With the American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Rhythm Society, Society of Cardiovascular Anesthesiologists, Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, and Society for Vascular Surgery. *J Am Coll Cardio* 2007;**50**(17):1707—32.
4 Group PS, Devereaux PJ, Yang H, Yusuf S, Guyatt G, Leslie K, et al. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): a randomised controlled trial. *Lancet* 2008;**371**(9627):1839—47.

5 Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;**283**(15):2008–12.

6 Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epid* 2009;**62**(10):1006–12.

7 Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions version 5.0.2. *The Cochrane Collaboration*; 2009 [updated September 2009].

8 Centre for Reviews and Dissemination. *Systematic reviews: CRD's guidance for undertaking reviews in health care [Internet]*. York: University of York [accessed 29.03.2010]. Available from: http://www.york.ac.uk/inst/crd/systematic_reviews_book.ht; 2009.

9 Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PMCochrane Diagnostic Test Accuracy Working Group. Systematic reviews of diagnostic test accuracy. *Ann Int Med* 2008;**149**(12):889–97.

10 Jeng JS, Liu HM, Tu YK. Carotid angioplasty with or without stenting versus carotid endarterectomy for carotid artery stenosis: a meta-analysis. *J Neurol Sci* 2008;**270**(1–2):40–7.

11 Ringleb PA, Chatellier G, Hacke W, Favre JP, Bartoli JM, Eckstein HH, et al. Safety of endovascular treatment of carotid artery stenosis compared with surgical treatment: a meta-analysis. *J Vasc Surg* 2008;**47**(2):350–5.

12 Ederle J, Featherstone RL, Brown MM. Percutaneous transluminal angioplasty and stenting for carotid artery stenosis. *Cochrane Database of Systematic Reviews* 2007;(4):CD000515.

13 Luebke T, Aleksic M, Brunkwall J. Meta-analysis of randomized trials comparing carotid endarterectomy and endovascular treatment. *Euro J Vasc & Endovasc Surg* 2007;**34**(4):470–9.

14 McManus RJ, Wilson S, Delaney BC, Fitzmaurice DA, Hyde CJ, Tobias RS, et al. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *BMJ* 1998;**317**(7172):1562–3.

15 Moher D, Fortin P, Jadad AR, Juni P, Klassen T, Le Lorier J, et al. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. *Lancet* 1996;**347**(8998):363–6 [see comment].

16 Gregoire G, Derderian F, Le Lorier J. Selecting the language of the publications included in a meta-analysis: is there a Tower of Babel bias? *J Clin Epid* 1995;**48**(1):159–63.

17 Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epid* 2002;**31**(1):115–23.

18 The Cochrane Collaboration [Internet]. Oxford [accessed 26.09.2009]. Available from: www.cochrane.org.

19 Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992;**305**(6857):786–8.

20 Borm GF, Donders AR. Updating meta-analyses leads to larger type I errors than publication bias. *J Clin Epid* 2009;**62**(8). 825.e10–830.e10.

21 Rerkasem K, Rothwell PM. Systematic review of the operative risks of carotid endarterectomy for recently symptomatic stenosis in relation to the timing of surgery. *Stroke* 2009;**40**(10): e564–72.

22 Culwell KR, Curtis KM, del Carmen Cravioto M. Safety of contraceptive method use among women with systemic lupus erythematosus: a systematic review. *Obst & Gynec* 2009;**114**(2 Pt 1):341–53.

23 Rebollo Aguirre AC, Ramos-Font C, Villegas Portero R, Cook GJ, Llamas Elvira JM, Tabares AR. 18F-fluorodeoxiglucose positron emission tomography for the evaluation of neoadjuvant therapy response in esophageal cancer: systematic review of the literature. *Ann Surg* 2009;**250**(2):247–54.

24 Deeks JJ. Issues in the selection of a summary statistic for meta-analysis of clinical trials with binary outcomes. *Stat Med* 2002;**21**(11):1575–600.

25 Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Medi* 2002;**21**(11):1539–58.

26 Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Method* 2008;**8**:79.

27 Higgins JP. Commentary: heterogeneity in meta-analysis should be expected and appropriately quantified. *Int J Epid* 2008;**37**(5):1158–60.

28 Patsopoulos NA, Evangelou E, Ioannidis JP. Sensitivity of between-study heterogeneity in meta-analysis: proposed metrics and empirical evaluation. *Int J Epid* 2008;**37**(5):1148–57.

29 Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**(8746):867–72.

30 Song F, Parekh-Bhurke S, Hooper L, Loke Y, Ryder J, Sutton A, et al. Extent of publication bias in different categories of research cohorts: a meta-analysis of empirical studies. *BMC Med Res Method* 2009;**9**(1):79.

31 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;**315**(7109):629–34.

32 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. *J Clin Epid* 2008;**61**(10):991–6.

33 Sutton AJ, Higgins JP. Recent developments in meta-analysis. *Stat Medi* 2008;**27**(5):625–50.

34 Lau J, Ioannidis JP, Terrin N, Schmid CH, Olkin I. The case of the misleading funnel plot. *BMJ* 2006;**333**(7568):597–600.

35 Agarwal R, Sinha AD. Cardiovascular protection with antihypertensive drugs in dialysis patients: systematic review and meta-analysis. *Hypertension* 2009;**53**(5):860–6.

36 Sanabria A, Dominguez LC, Valdivieso E, Gomez G. Prophylactic antibiotics for mesh inguinal hernioplasty: a meta-analysis. *Annals of Surgery* 2007;**245**(3):392–6.

37 Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biome* 2000;**56**(2):455–63.

38 Moreno SG, Sutton AJ, Ades AE, Stanley TD, Abrams KR, Peters JL, et al. Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Med Res Method* 2009;**9**:2.

39 Bown MJ, Bolia A, Sutton AJ. Subintimal angioplasty: meta-analytical evidence of clinical utility. *Euro J Vasc & Endovasc Surg* 2009;**38**(3):323–37.

40 Yang SH, Yang KH, Li YP, Zhang YC, He XD, Song AL, et al. Breast conservation therapy for stage I or stage II breast cancer: a meta-analysis of randomized controlled trials. *Ann Oncol* 2008;**19**(6):1039–44.

41 Bax L, Yu LM, Ikeda N, Moons KG. A systematic comparison of software dedicated to meta-analysis of causal studies. *BMC Medical Research Methodology* 2007;**7**:40.

42 STATA [Internet] [accessed 13.05.2010]. Available from: http://www.stata.com.

43 The R project [Internet] [accessed 13.05.2010]. Available from: http://www.r-project.org/.

44 Comprehensive meta-analysis [Internet] [accessed 12.07.2010]. Available from: http://www.meta-analysis.com.

45 Bax L. *MIX 2.0. Professional software for meta-analysis in excel. Version (2.0.0.7)*. BiostatXL, http://www.meta-analysis-made-easy.com; 2010.

46 Statsdirect [Internet] [accessed 13.05.2010]. Available from: http://www.statsdirect.com/.