🖹 ARTICLE

Meta-analysis, collaborative overview, systematic review: what does it all mean?

Mike Clarke | 9 min read

Summary

Bringing together the evidence from randomised controlled trials is increasingly common. Depending on the techniques used, the process might be called a systematic review, meta-analysis or collaborative overview. All relevant trials are identified and those of a satisfactory standard are reviewed. In a meta-analysis, each trial is assessed separately and the summary statistics are then combined to give an overall result. Neither meta-analyses nor systematic reviews are a substitute for prospective clinical trials, rather, they are a complement to them.

Introduction

In many areas of health care, there may be modest differences between treatments. However, these differences are worth knowing about, especially if one treatment is relatively easy to give and the condition it is used for is common. Moderate differences can be detected reliably through adequately-sized trials which accrue new patients, the analysis of past trials, or a combination of both.

Whichever approach is followed, systematic biases and chance effects must be minimised. Systematic bias within trials is controlled mainly by ensuring that the method of randomisation guarantees that the next treatment to be allocated cannot be predicted before the patient is entered onto the trial, and by analysing the trial using the intention-to-treat principle. The total number of events must be high enough so that any chance excess will be smaller than the excess that would be seen if there was a genuine difference between the treatments.[1] Therefore, chance effects are minimised by increasing the number of patients studied.

Randomised controlled trials or systematic reviews?

Some prospective reviews are under way in which it has been agreed that ongoing trials will pool their results[2], but systematic reviews are usually based on trials from the past. This has both advantages and disadvantages. If many trials have assessed a therapeutic question, bringing them together should give a relatively quick answer on the effects of the treatments. If prolonged follow-up is important, trials from the past might provide this information immediately, whereas prospective trials may not produce results for years. Conversely, insufficient information might be available from past trials and sufficient randomised evidence will need to be obtained prospectively. However, without a systematic review to identify prior trials, a truly informed decision cannot be made on whether sufficient information already exists.

An advantage of a single randomised trial over a review is that the therapeutic question will be addressed in the same way in all patients. The statistical techniques used for meta-analyses control for differences that arise between trials, but the existence of these differences may mean that a single trial is perceived, rightly or wrongly, as more reliable.

Systematic reviews, meta-analyses and collaborative overviews

The terminology used to describe the combination of trials varies, but the above terms are widely used.

Systematic reviews

This is the broadest term and describes reviews where the design and methods are defined (including inclusion and exclusion criteria, and how trials will be identified), and preferably written into a review protocol before the review commences. A systematic review seeks to include all relevant trials. The key principle for any systematic review is that as much as possible of the randomised evidence is identified and included in any analyses that are done. This differs from the more 'traditional' form of review in which only a selected group of trials, which is probably biased in some way, is considered. Including all trials minimises systematic biases that might arise if just a subset is used. Increasing the amount of data minimises chance effects.[3]

Meta-analysis

A systematic review does not necessarily contain a statistical synthesis of the results from the included trials. For example, the reviewer might conclude that their design makes the identified trials too dissimilar to be combined or that the results available for each trial may not be able to be combined. A review of the use of surgery for stress incontinence concluded that, although more than 30 prospective studies had been identified, these could not be combined because of the wide variation in the design and outcome assessment of the studies.[4]

If the results of the various trials are combined, this is a meta-analysis. To be systematic, this still needs to attempt to include all relevant trials since a meta-analysis would be non-systematic if, for example, the results were combined for only a small number of the trials relevant to a review. Some authors now seek to clarify this by using both terms to describe a meta-analysis which comes from a systematic attempt to find all relevant trials.[5]

Collaborative overviews

The data to be included in a meta-analysis can be extracted from each trial publication or obtained directly from the trialists. A collaborative overview is a refinement of this. It involves the central collection of data on each and every randomised patient and is also known as a meta-analysis based on individual patient data.

The Early Breast Cancer Trialists' Collaborative Group overview (EBCTCG) is an example of such a project. The Group tries to identify all randomised trials of the treatment of breast cancer and some data are sought for each of the women entered into these trials. In 1992, data were brought together on 75 000 women to provide reliable answers on the effects of tamoxifen, chemotherapy and immunotherapy.[6] This project takes place every 5 years so that updated data can be collected for all women and all trials. The results of the overviews will be published and also made available through the Breast Cancer Collaborative Review Group of the Cochrane Collaboration.[7]

Collaborative overviews of individual patient data have additional advantages over other types of systematic review. Generally, the most important is that time-to-event calculations can be done. In reviews with death as an outcome, these calculations may show a prolongation of survival for one treatment group. This might not be seen from follow-up data for a fixed time point. In addition, it might be easier for a trialist to send individual patient, rather than aggregate, data and it will be easier for a small amount of extra information to be supplied. For example, if further follow-up becomes available on some patients, the trialist can simply send these details instead of preparing new tables. Finally, central checking of the patient data may identify problems which can be rectified through consultation with the trialist.[8]

Which trials should be included

One criticism that is raised against systematic reviews is that they compare apples and oranges. The quick response is that this is appropriate if one is interested in fruit, but not if one is interested only in apples or oranges! There is the problem of heterogeneity resulting from the basic differences that exist between trials. This includes differences in areas such as the quality of their design, the eligibility criteria or the treatments used. The question of which trials to combine is very important. Combining trials unnecessarily may increase the statistical precision of the result, but it decreases its clinical reliability if the trials are too dissimilar. Trials should be combined only if their designs are sufficiently alike and this decision must not be based on the trial's result. If the 'lumping' together of different trials is controversial, presenting the review so that the contribution of each trial can be seen allows the reader to 'split' the review and recalculate the overall result accordingly.

Fig. 1 shows how the representation of the data for each trial in a meta-analysis allows this. The data are from the EBCTCG overview of ovarian ablation in breast cancer[9] and show a 18.4% reduction in the annual odds of death for women allocated to ovarian ablation rather than to the control group if all of the trials are considered. This overall result is obtained by summing the statistics (O-E and variance) for each of the trials. Thus, if, for some reason, the user of a review wished to calculate the result without a particular trial, or just for a subgroup of the trials, they could do so from these individual statistics. The O-E statistic for each trial is calculated from the number of deaths 'observed' (O) in the treatment group minus the number of deaths that would have been 'expected' (E) if the deaths had been evenly distributed between treatment and control patients. If O-E is negative, this favours treatment as it means that fewer patients died in that group. The variance is calculated from the O-E statistic and the total number of deaths in the trial. The larger the variance, the larger the statistical weight of the trial in the meta-analysis.

Trial identification
The most important step in a systematic review is trial identification. Every attempt should be made to ensure that as many as possible of the potentially eligible trials are found. This is especially important because of publication bias: trials with positive results are more likely to be published. As a consequence, the results of published and unpublished trials could be systematically different. Unless all trials are sought, regardless of their publication status, the review may contain a biased set of studies. Then, regardless of how the data are handled, a meta-analysis may be mathematically accurate, but clinically unreliable.

Collection and analysis of data
If too few trials are identified for their results to be combined, the systematic review still makes an important contribution by highlighting the lack of sufficient evidence. This should be a stimulus for an adequately-sized randomised trial. If there are sufficient trials for a meta-analysis, the principles of decreasing systematic biases and chance effects must be applied when data are collected. All relevant trials should be included. If this is not possible, any trials which do not contribute data must not be so numerous or unrepresentative as to affect materially the result of the meta-analysis.

Fig. 1

An overview of randomised trials of ovarian ablation versus control for women with early breast cancer.

Each trial is described by one line of data showing the trial name, the number of deaths and women in the ovarian ablation and the control groups, and the statistical results of the analyses within that trial. The area of the black square is proportional to the amount of information contributed by each trial and the horizontal line through it represents the 99% confidence interval associated with the result of the trial. The diamonds represent the statistical subtotals and totals for the average effects within the trials and their widths represent the associated 95% confidence intervals.

(Included with permission of © The Lancet Ltd from Early Breast Cancer Trialists' Collaborative Group. Ovarian ablation in early breast cancer: overview of the randomised trials. Lancet 1996;348:1189-96.)

/upload/issue_files/2004_232_1.jpg

The ultimate aim is that all randomised patients, and no non-randomised patients, from all relevant trials are included and that they are analysed using the intention-to-treat principle. If a trial has not been published, the reviewer needs to contact the trialist for the necessary data. Even if a trial has been published, it will often be necessary to obtain additional information from the trialist. For example, the outcomes needed for the meta-analysis might not have been reported adequately or the published results may not follow the intention-to-treat principle. Collection of data from trialists will also allow them to supply more up-to-date and complete data than were published. The fact that all published results are 'frozen in time' will be important if further follow-up could change the results of the trial or the meta-analysis. The data collected for the meta-analysis can be either aggregate or individual patient data. Contacting the trialist may also help clarify the design of their trial. Occasionally, this may lead to it being reclassified as ineligible because it was not properly randomised or the treatment comparison had been misunderstood.

Subgroup analyses

Dividing results into different types of patients and outcomes requires cautious interpretation. If these analyses are to be done, common definitions should, where possible, be used for all trials. This may be easier with individual patient data. However, if the overall result is subdivided in any way, the reviewer and reader need to be wary. The more subgroup analyses are performed, the more likely it is that a statistically significant, but incorrect, result will be found due purely to the play of chance. Any subgroup analysis should best be regarded as generating a hypothesis for testing in the future. It is often more reliable to assume that the overall result is as good an estimate (if not a better one) for a particular group of patients than that obtained by looking at these patients within the meta-analysis.

A good illustration of this comes from an investigation of the combined effects of chance, subgroup analysis and publication bias.[1] Forty-four randomised trials were simulated by rolling different coloured dice, with each roll of the dice yielding the outcome for one 'patient'. Each participant simulated two trials and it was pre-specified that subgroup analyses would be done which distinguished each participant's first trial from their second trial. Overall, chance alone produced a finding that 'treatment' was non-significantly better than 'control' with a reduction in the odds of death of 11% ± 11. However, when the pre-specified subgroup analysis was done, it was found that (again due to chance alone) all of the benefit came from the second trials. The first trials produced exactly the same total number of deaths in each group of 'patients': a 0% ± 15 reduction in the odds of death. The second trials showed a benefit which, although not significant, was quite large: 24% ± 15. The point estimate for this would be equivalent to 38 lives saved per thousand patients treated.

This experiment therefore concluded that the premise 'Don't ignore chance effects' (DICE) should always be remembered, particularly in regard to the results of subgroup analyses.

Statistics

If a meta-analysis is done, each trial is analysed separately and the overall result comes from combining these summary statistics. No assumptions are made about patients in one trial being directly comparable with those in another, or about treatment differences being the same in each trial.

Conclusion

These methods of combining and reviewing data are an important means for assessing health care interventions. They are a complement to, not a replacement for, prospective randomised controlled trials. The reviews must be conducted reliably. This is achieved by minimising systematic biases and chance effects. The maximum amount of randomised evidence needs to be gathered. The most important part of the process is that all relevant trials are identified. This will help to obtain statistically accurate results. The next step is interpreting the clinical relevance of these results.

References

1. Counsell CE, Clarke MJ, Slattery J, Sandercock PA. The miracle of DICE therapy for acute stroke: fact or fictional product of subgroup analysis? Br Med J 1994;309:1677-81.

2. Cholesterol Treatment Trialists' (CTT) Collaboration. Protocol for a prospective collaborative overview of all current and planned randomized trials of cholesterol treatment regimens. Am J Cardiol 1995;75:1130-4.

3. Clarke MJ, Stewart LA. Obtaining data from randomised controlled trials: how much do we need for reliable and informative meta-analyses? Br Med J 1994;309:1007-10.

4. Black NA, Downs SH. The effectiveness of surgery for stress incontinence in women: a systematic review. Br J Urol 1996;78:497-510.

5. Aker PD, Gross AR, Goldsmith CH, Peloso P. Conservative management of mechanical neck pain: systematic overview and meta-analysis. Br Med J 1996;313:1291-6.

6. Early Breast Cancer Trialists' Collaborative Group. Systematic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy. 133 randomised trials involving 31 000 recurrences and 24 000 deaths among 75 000 women. Lancet 1992;339:1-15, 71-85.

7. Breast Cancer Collaborative Review Group. In: The Cochrane Library [database on disk and CDROM]. The Cochrane Collaboration. Oxford: Update Software, 1997-. Updated quarterly.

8. Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. Stat Med 1995;14:2057-79.

9. Early Breast Cancer Trialists' Collaborative Group. Ovarian ablation in early breast cancer: overview of the randomised trials. Lancet 1996;348:1189-96.