How to cite this document: Ryan R; Cochrane Consumers and Communication Review Group. 'Cochrane Consumers and Communication Group: meta-analysis. http://cccrg.cochrane.org, December 2016 (accessed DATE).

Cochrane Consumers and Communication Group reviews:
Meta-analysis

- See the *Cochrane Handbook for Systematic Reviews of Interventions*, Chapters 9, 11 and 12.

For many Cochrane Consumers and Communication (CCC) reviews, narrative (descriptive) synthesis is the only appropriate approach to data synthesis. Narrative synthesis will therefore always be part of the results section. If authors also decide to perform meta-analysis, it needs to be clearly justified.

Whether the data is analysed using meta-analysis or through narrative synthesis, the analysis should consider each of the following questions:

- What is the direction of effect (positive, negative, unclear)?
- What is the size of the effect?
- Is the effect consistent across studies?
- What is the strength of evidence (quality or certainty) for the effect?

Unless the meta-analysis is properly planned, and appropriate both to the review question and to the characteristics of the included studies, it may not produce useful or meaningful results. Worse, if poorly planned or executed, meta-analysis may produce seriously misleading results.

What is meta-analysis?

- *See the Cochrane Handbook,* Chapter 9, especially section 9.1.

Meta-analysis is a statistical method for pooling the results of several studies reporting the same outcome, in order to gain a better estimate of the effect size of an intervention. By converting the outcomes of different studies to a common measurement such as effect size, then means and proportions can be averaged across studies. If there are several studies on the effect of an intervention with varying directions or varying significance of outcome, combining these studies via meta- analysis may aid in making a judgement about the usefulness of the intervention.

Is meta-analysis appropriate?

Meta-analysis is appropriate when the following criteria can be met by the studies of interest:

- All the outcomes are comparable and can be pooled meaningfully.
- All the interventions and comparators are the same, or at least similar enough to be combined meaningfully.
- The correct data are available for the included studies; for example, means and standard deviations are necessary to analyse a continuous outcome (Note: even if the results are not reported in this format by the study authors, mean and standard deviation can often be calculated from the data provided).

Two studies is a sufficient number to perform a meta-analysis, provided that those two studies can be meaningfully pooled and provided their results are sufficiently 'similar'.

Whether studies are considered sufficiently 'similar' can be interpreted using a measurement of 'heterogeneity,' the extent of which guides the meta-analysis. Too much heterogeneity indicates that meta-analysis may be contra-indicated (see CCC 'Heterogeneity and subgroup analysis' guide at http://cccrg.cochrane.org/author-resources).

Please also note that if continuous data are skewed it may not be appropriate to summarise these using means and standard deviations, especially when the study size is small. There are different ways that such results can be summarized and presented, see the Handbook, section 9.4.5.3.

Conducting the meta-analysis

The first stage in conducting a meta-analysis is choosing the appropriate outcome measure. This will come from the review question and from author knowledge of the literature in the area. The outcome measure needs to be something that is interpretable and likely to be reported. Often this will be a dichotomous (or binary, yes/no) variable, such as patient death, recovery or behaviour. In other cases, the outcome will be a continuous measure, for example, change in blood pressure, time to recovery, knowledge or satisfaction.

When the outcome is continuous, the appropriate measure for meta-analysis is often the 'mean difference' (MD) statistic. Odds ratios, risk ratios or risk differences are used for dichotomous outcomes.

- *See the Cochrane Handbook,* Chapter 9, especially section 9.2.

Weighting of studies

Meta-analysis works by averaging the effect sizes (eg. mean difference, log odds ratio, risk difference) of the included studies. The average must account for the different amount of

information in each study and this is achieved using weights. The weights are calculated using the sample size and the variability within each study. The best estimate across all studies is the weighted average of the effect sizes. RevMan calculates the weights and the weighted average automatically.

Dealing with heterogeneity

- See the *Cochrane Handbook,* Section 9, especially sections 9.5 and9.6.
- See also the CCC 'Heterogeneity and subgroup analysis' guidance at http://cccrg.cochrane.org/author-resources.

Large differences between study results will be indicated by the heterogeneity statistic (I^2), which quantifies the degree of heterogeneity or inconsistency.

The I^2 statistic can be interpreted roughly as:

- 0% to 40%: might not be an important level of inconsistency
- 30% to 60%: may represent moderate heterogeneity*
- 50% to 90%: may represent substantial heterogeneity*
- 75% to 100%: considerable heterogeneity*

*The importance of the I^2 value depends on (i) the size and direction of effects and (ii) the strength of evidence for heterogeneity (e.g. P value from the Chi^2 test; adapted from the *Cochrane Handbook* section 9.5.2).

When the heterogeneity statistic is very large (i.e. there is a lot of variability in the results), authors might choose to abandon the meta-analysis. It is not necessary to include meta-analyses in a systematic review, and where heterogeneity is high it may be more appropriate to analyse the results narratively, rather than statistically.

However, there is a way to cope with moderate heterogeneity – using a random-effects model for the meta-analysis. The random-effects model incorporates the differences between studies in the calculations and (usually) increases the width of the confidence interval around the pooled estimate of effect, so giving a more conservative estimate of effect. The alternative to the random-effects model is the fixed-effect model which assumes that all studies are consistent and similar.

Deciding whether to use a fixed-effect model or a random-effects model is something of a judgement call. The test for heterogeneity can help make the decision, but it not always very helpful. If the random- effects model and the fixed-effect model produce substantially different pooled estimates then this is an excellent indication of heterogeneity and the random-effects model is the preferred model. If the two models yield similar pooled estimates then the fixed-effect model is preferred, because usually it will have a narrower confidence interval; that is, it is more precise than the random-effects model.

In reviews of complex interventions, where there is often inherent variability in the design, delivery and other features of the interventions, it is often most appropriate to use the random effects model for meta-analysis.

Sensitivity analysis

- See the *Cochrane Handbook*, section9, especially section 9.7.

Sensitivity analysis is one way to explore heterogeneity in a meta-analysis and is often used in Cochrane reviews. Simply put, one or more studies are removed from the meta-analysis to examine the effect removal of the study or studies has on the pooled intervention effect. A study of lower quality or one that appears to be an outlier are examples of the reasons for removing a study in a sensitivity analysis. There is no fixed rule to judge whether a study's removal has 'significantly' influenced the pooled treatment effect; this judgement can only be made by assessing the clinical or public health importance of the change in effect size.

Subgroup analysis

- See the *Cochrane Handbook*, Section 9, especially section 9.6
- See also the CCC 'Heterogeneity and subgroup analysis' guidance at http://cccrg.cochrane.org/author-resources.

Subgroup analysis is similar to sensitivity analysis in that it examines the effect of certain studies on the pooled treatment effect. However, rather than examining an element such as study quality, it examines a particular aspect of the study delivery. In subgroup analysis, particular factors that there is good reason to think might influence the effects of the intervention are identified and used to 'divide up' the studies of interest. For example, for one outcome measure there may be two subgroup analyses about the way the intervention was delivered (that is, a mail intervention versus a telephone intervention) or the way the effect was measured (that is, using two different measurement scales), or there may be features of the population receiving the intervention (that is, the effects in adults versus children), the setting (studies in primary care versus hospital), or other factors.

Subgroup analysis should be kept to a minimum, and pre-specified and justified at the protocol stage of the review. The planned analyses should be followed at review stage (if sufficient data are available) to minimise selective reporting or over-interpretation of the results based on findings.

Meta-regression

- See the *Cochrane Handbook,* Section 9, especially section 9.6.4 to 9.6.6.

Meta-regression is an extension of a random-effects meta-analysis and can be used to explore heterogeneity in greater detail. It estimates the effect that one or more study

variables have on the effect size. These may be variables such as year of study, dose of drug or geographical location.

Meta-regression is not often performed in Cochrane reviews, and is usually not recommended unless there are at least 10 studies included in a meta-analysis. If there is a strong reason to include meta-regression in your review you will need to seek the assistance of a statistician.

Publication bias

- See the *Cochrane Handbook*, Chapter 10, for information on reporting biases, including publication bias.

Publication bias may happen because studies with non-significant results are less likely to be published than those with positive findings. This means that a meta-analysis of published studies may overestimate the true effect size. Funnel plots, available through RevMan, can give some indication of the possibility of publication bias, though it should be remembered that an asymmetrical funnel plot may also result from other issues such as study quality. Though there are methods available to correct for publication bias the main thing is to ensure that the review authors have raised the issue, so that readers are aware of the possibility.

Tips for authors undertaking analysis

- See chapter 12 of the *Cochrane Handbook*, especially Sections 12.4 to 12.7.

- Other resources for review authors on meta-analysis:
    - http://www.metaanalysis.com/pages/why_do.html
    - http://www.statsdirect.com/
    - http://www.bmj.com/cgi/content/full/315/7121/1533
    - http://www.pitt.edu/~super1/lecture/lec1171/index.htm

Below are general prompts to help authors with writing and organising the results and analysis. Authors should also consult the *Cochrane Handbook* for advice and to ensure that reviews meet the guidelines for Cochrane reviews.

Selecting outcomes

- There should be clear correspondence between the objectives of the review, the interventions and the outcomes selected for the review.
- In general, primary or direct outcomes are best; surrogate outcomes may be less reliable in terms of what they actually measure/ reflect.
- We recommend that review authors do not just select the primary outcome of each included study and combine these (statistically or narratively). With complex interventions there are often many outcomes, measured in many different ways, and

lumping them together may make them less meaningful than considering them individually or in groups of outcomes that have a theoretical basis or clear logic.

The outcome measures chosen

- If outcomes are to be combined statistically, it is important to be clear about what the outcomes mean; what a change in them means for the effects of the intervention; and how these should be interpreted. For example, an outcome such as a high score on a symptom checklist may represent an undesirable outcome, and one that would ideally be decreased by a particular intervention (eg a self-management intervention). On the other hand, an outcome such as a high score on the number of symptom-free days may represent a desirable outcome.
- When examining your results and the outcomes reported by individual studies, it is vital to understand the practical and/or clinical implications of the outcomes, and what a change in each outcome really represents.
- It may also be necessary to transform the results into a common language, for the sake of clarity. In the example above, it may make the most sense to readers to talk about changes in symptoms associated with a particular intervention: transforming the results expressed as an increase in symptom-free days to a decrease in symptoms, for example, might be clearer to readers.

Do the results make sense?

- If the results of pooled analyses are very unexpected or run opposite to expectations from previous studies, re-check the data extracted from included studies, in case errors have been made in data extraction or entry: these are very common.
- Likewise, if individual studies conclude that the intervention effect is in one direction, while a synthesis suggests an effect in the opposite direction, you should be suspicious of the results and check them again (eg. outcomes may have been entered incorrectly (ie the wrong way around); the outcomes may not have been adjusted correctly to account for differences in scale (positive or negative); and so on. The *Handbook* recommends against using change scores for this reason, among others: use endpoint scores if possible.
- It can be useful to extract the text of the conclusions of the included studies as part of the data extraction process to enable internal cross-checking of the results data when writing the review.

Presenting the results consistently

- See the *Cochrane Handbook*, Chapter 11, especially section 11.7.

- See also the additional CCCG resources, available at http://cccrg.cochrane.org/author-resources, in particular:
    - 'Identifying comparisons in CCCG reviews'
    - 'Describing results'

- It can be challenging to organise the results of a review, especially if there are a large number of included studies with a range of outcomes, and if both narrative synthesis

and meta-analysis have been used.

- It is very important that the results are presented in a clear and consistent way throughout the review, and some planning can help to make this process easier.
- As a general rule, if meta-analysis and narrative synthesis are both to be reported in the review, it is essential that they correspond exactly and follow a logical sequence.

  Authors should use the following as a general guide:
  - The results section should be organised to follow the order of the comparisons and outcomes as specified in the protocol and earlier in the review (Objectives and Selection criteria).
  - Presentation and reporting of the results of meta-analysis must follow the same structure as the narrative synthesis (ie they should be integrated within the review text).
  - If forest plots or other graphs or figures are included in the review, they must be presented logically and/or referred to in the review text, along with the numerical effect estimate and a measure of variability (eg. 95% confidence interval). Citing P values for statistically significant results (preferably the exact P value) is also desirable in the text; significant degrees of heterogeneity should also be reported in the text if detected in the meta-analysis for particular outcomes.
  - The numbers cited in the review text and those in the graphs/ figures must correspond exactly. If different statistical models are to be used (eg. random-effects versus fixed-effect models) and different effect estimates arise as a result of the different methods used, this must be made clear in the text.
  - Where numerical results are referred to in the text, their meaning should be clear (eg. increase/decrease in the outcome being assessed) and they must also refer to the appropriate graph/figure number if this has been included in the analysis.
  - The description of the results in the text should not only consider statistical significance of the results: instead the findings should be described in terms of the direction, size and variability of the effects, and the quality of the evidence for that outcome. For more information on this, please see the following resources, available at http://cccrg.cochrane.org/author-resources:
    - 'How to GRADE'
    - 'Describing results'
  -

Last updated: 1st December 2016

We are grateful for the expert input of Associate Professor Damien Jolley and Ms Kelly Allen (Monash University) to the development of this advice