Statistical issues and common analysis queries

Cochrane Heart - Methods Guidance

Updated October 2019 with references to Version 6 of the Cochrane Handbook for Systematic Reviews of Interventions, available from: https://training.cochrane.org/handbook.

1. Heterogeneity

Please don't use a simple threshold (e.g. $I^2 > 50\%$) to diagnose heterogeneity. It is important to consider the direction and magnitude of effect and the degree of overlap between confidence intervals on the forest plot. There is substantial uncertainty in the value of I^2 when there is only a small number of studies; please also consider the P value from the Chi^2 test.

See MECIR standard C63 and Section 10.10.2 of the *Cochrane Handbook* (v.6).

2. Presentation of Subgroups in Data and Analyses tables

The problem: The Revman tutorial shows you how to set up a single subgroup analysis, but not how to deal with multiple subgroups. You can't introduce a second subgroup to the same forest plot, as it invalidates the test for subgroup differences, as the same people may be appearing in more than one subgroup. It is tempting to use the top level heading (1. Comparison) to describe each subgroup analysis, but this goes against **MECIR standard R75**, which states that appropriate use of the hierarchy is mandatory.

Cochrane Heart's preferred approach – start with the main analysis of all trials, then start a new outcome for the first subgroup. In this example, the main comparison is drug A vs drug B, with 2 outcomes of interest (mortality and readmissions) and two subgroup analyses (children vs. adults, and low vs high dose of drug A). It would be helpful to use capital letters to indicate the main analysis for each outcome.

1. Drug A vs Drug B
    - 1.1 MORTALITY
    - 1.2 Mortality – by age
        - 1.2.1 Studies in children
        - 1.2.2 Studies in adults
    - 1.3 Mortality – by dose
        - 1.3.1 High dose drug A
        - 1.3.2 Low dose drug A
    - 1.4 READMISSIONS
    - 1.5 Readmissions by age
        - 1.5.1 Studies in children
        - 1.5.2 Studies in adults
    - 1.6 Readmissions – subgroup analysis by dose
        - 1.6.1 High dose drug A
        - 1.6.2 Low dose drug A

3. Subgroup analysis

Interpretation of subgroup analysis should be based on the formal statistical test for differences between groups. Please do not compare the P values of different groups within a subgroup analysis and conclude that there is an effect in group A but not groups B and C. Consider the degree of overlap between confidence intervals, and bear in mind that there may be a lack of information for some groups rather than an absence of effect.
See MECIR standard C68 and Section 10.11.5.2 of the *Cochrane Handbook* (v.6).

Authors are encouraged to use the RoB2 tool for new reviews. Chapter 7 of the *Cochrane Handbook* (v.6) describes the recommended approach for considering bias and conflicts of interest. Please note that the TACIT tool for assessing the influence of conflicts of interest is currently being developed (http://tacit.one).

If you are still using the original Cochrane 'Risk of bias' method, please note Points 4 and 5 below.

4. Overall risk of Bias

If you are doing a sensitivity analysis and only want to include studies at low risk of bias, please consider which domains are of key importance for each outcome. Cochrane Heart will only accept the following domains for categorising studies' overall risk of bias for a sensitivity analysis:

1. Random sequence generation.
2. Allocation concealment.
3. Incomplete outcome data.
4. Selective outcome reporting.
5. Blinding *where this is relevant to the outcome*, *considered separately for subjective and objective outcomes.*

In particular, Cochrane Heart will not accept industry funding as a reason for marking studies at high risk of bias. Any clear undue influence by funders should be justified as part of the assessment of bias in other domains. See also Point 5 - Industry Funding.
See MECIR standards C59 and C60 and Sections 7.6.1 and 7.8.6 of the *Cochrane Handbook* (v.6).

5. Industry funding

Industry funding should not in itself be used as a reason for downgrading GRADE ratings. Any influence of funders should be taken into consideration when assessing the relevant risk of bias domains, which may then feed into the GRADE assessment. Details of funders and any influence they had should be provided in the Characteristics of Included Studies table.
This is mentioned in Section 8.15.1.5 of the *Cochrane Handbook* (v.5.1; 2011).

6. Publication bias

If you have at least 10 studies in the meta-analysis, we encourage you to perform a formal statistical test for funnel plot asymmetry. If you have a small number of studies, please be aware that the ability to detect publication bias is largely diminished, so it is difficult to exclude the presence of publication bias. Please note that funnel plots are a general way of assessing whether small studies show different effect estimates than larger studies (small-study effects). Non-reporting bias is only one component of this, so authors should consider whether there may be other causes of any observed asymmetry.

See MECIR standard C73 and relevant sections of the *Cochrane Handbook* (v.6): Section 13.3.5.2 for advice on funnel plots, Section 13.3.5.3 'Different reasons for funnel plot asymmetry', and Section 13.3.5.4 for information about tests for funnel plot asymmetry.

7. Inclusion of zero event studies in a meta-analysis and forest plot

For studies with zero events in one arm but one or more events in the other arm, Revman adds a continuity correction and can include them in the meta-analysis.

The problem: If there are zero events in both arms, Revman can't include this study in the meta-analysis. By default, it excludes the study from the pooled estimate, but retains it in the forest plot and total number of patients.

- This has the advantage of showing clearly on the forest plot the overall number of events and participants in all relevant studies, including those with zero events.
- But by including the study in the 'total' for the analysis, it makes the meta-analysis look like it is based on more studies and participants than it actually is. This automatically goes through to the Summary of Findings table, automatically linked analyses etc.

Cochrane Heart's preferred approach: We ask authors to remove any studies with zero events in both arms from the forest plot and discuss such studies narratively alongside the meta-analysis results. A brief statement can also be added to the Comments column of the Summary of Findings table. It would be helpful to state what proportion of the total number of participants and studies contributed to the meta-analysis.

An exception to this is where you have a multi-arm study that can be included in the overall analysis, but splitting into subgroups gives zero events in one arm of the study. In this case, it would be preferable to include the zero events arm so that the same number of people are included in the main analysis and the subgroup analyses.

Please see Section 10.4.4.2 of the *Cochrane Handbook* (v.6).

8. Dealing with multi-arm studies

If you have a multi-arm study with more than two relevant arms, please add this only once to the Table of Included studies. This avoids 'double counting' the study in terms of the number of RCTs you have, and avoids the study appearing twice in the Risk of Bias diagrams. If there are differences in risk of bias between arms (for example for blinding or attrition), this can be added as a comment in the "support for judgement" box.

As Revman will not allow you to amend the name of the study, you will then need to use footnotes to distinguish between arms in the forest plots. If both arms are appearing in the same forest plot, please ensure that you have split the control group in half (or taken other appropriate steps to avoid double counting).

9. Imputation of missing data

If a small proportion of your studies have missing data and it is reasonable to impute any values, please ensure that the process is clearly described in your Methods section. It would also be appropriate to assess the impact of changing the assumptions used by the imputation in sensitivity analyses.

See MECIR standard R44 and Chapter 6 of the *Cochrane Handbook* (v.6).

10. Composite outcomes

It is preferable to avoid using composite outcomes (such as Serious Adverse Events) and use individual outcome measures as far as possible. If a review uses a strict definition of SAE, and the included study uses exactly the same definition, the SAE data should be reliable. However, problems arise when review authors take individual events (such as stroke, myocardial infarction) and add them together to 'build' their own SAE data. This risks double counting people who may have had more than one such event, and also risks missing other events which occurred but were not reported by trial authors (i.e. they didn't present data for myocardial infarction, but you can't assume there weren't any when summarising SAE).

See Chapter 19 of the *Cochrane Handbook* (v.6).

11. Interpretation of results

Many authors find it helpful to use the document 'describing results' kindly produced by the Cochrane Consumers and Communication Group,[1] and available from: heart.cochrane.org/for-authors/essential-resources. If you use this table to guide your interpretation of results, please ensure that you are consistent with the phrasing in all areas of the review.

The phrasing refers to 'little or no difference' in outcome for results showing no important benefit/harm or a null effect. For Heart reviews, it is preferable to make a distinction between 'no effect' and 'little effect' if the confidence intervals are sufficiently tight around the null point.

[1] Ryan R, Synnot A, Hill S (2016) Describing results. Cochrane Consumers and Communication Group, available at: http://cccrg.cochrane.org/author-resources. Version 2.0 December 2016.

12. Statistical significance

It is important to avoid describing results as statistically significant (or not), and instead interpret the size and importance of the effect. Authors should carefully consider what constitutes an important effect for their outcomes of interest.

Whilst the width of confidence intervals is considered as part of the GRADE assessment of imprecision, please note the following:
*"If the confidence interval for the estimate of the difference in the effects of the interventions overlaps with no effect, the analysis is compatible with both a true beneficial effect and a true harmful effect. If one of the possibilities is mentioned in the conclusion, the other possibility should be mentioned as well." (**Section 15.6.4*** **of the *Cochrane Handbook*, v.6).**

For example, if an 8% difference is seen as an important clinical effect, the following could be considered as showing a benefit: RR=0.92 (CI 0.83, 1.01). Whilst the confidence interval includes the null (so is theoretically compatible with there being no effect, or perhaps a lack of evidence for an effect depending on sample size), it does not include an important harmful effect using the example of 8%.

By contrast, RR=0.90 (CI 0.72, 1.13) would suggest a lack of evidence for an effect – whilst the point estimate suggests an important benefit, the confidence interval is wide enough to also include an important harm.