Statistical issues and common analysis queries

Cochrane Heart Group

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 effects 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 9.5.2 of the Cochrane Handbook (v5.2)[1].

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

[1] Deeks JJ, Higgins JPT, Altman DG (editors) on behalf of the Cochrane Statistical Methods Group. Chapter 9: Analysing data and undertaking meta analyses. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

3. Subgroup analysis

Interpretation of subgroup analysis should be based on the formal statistical test for differences between groups. Please don't 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 Cochrane Handbook v.5.2 Section 9.6.3.1[1]

4. Sensitivity analysis

Whilst subgroup analysis looks at whether the effect is different in particular groups of people (or studies conducted in different groups of people), sensitivity analysis aims to test the assumptions used in the statistical analysis. This involves repeating the analysis but restricting to certain studies, for example those of higher quality.

Unlike subgroup analysis, there is no formal statistical test that can be used for sensitivity analysis, so authors must make informal comparisons between the different ways of estimating the effect under different assumptions. Please don't judge whether there is a difference between the main analysis and sensitivity analysis by comparing changes in p values.

See MECIR standard C71 and Cochrane Handbook v.5.2 Section 9.7[1]

5. 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 8 -Industry Funding.

See MECIR standards C56 and C58 and Cochrane Handbook v.5.2 Sections 8.7 and 8.12.2[2]

[2] Higgins JPT, Altman DG, Sterne JAC (editors). Chapter 8: Assessing risk of bias in included studies. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

6. Publication bias

Please note that the updated version of Chapter 10 (Handbook v5.2)[3] states that the minimum number of trials required to generate an adequately powered funnel plot may be considerably higher than 10 in some circumstances (Section 10.4.3.1). There is also a new section 10.4.4.7 that discusses regression methods for dealing with potential reporting biases. If you have sufficient studies for a funnel plot, we would encourage you to perform a formal statistical test for 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.

7. GRADE assessments

Please see Handbook Section 11.2.2 of the updated version of the Cochrane Handbook (v5.2)[4] for a detailed description of making GRADE assessments and downgrading evidence. Table 11.3.a in v.5.2 of the Handbook provides an excellent framework for justifying downgrading decisions in footnotes.

- Risk of Bias - only assessments in *relevant domains* and for *relevant studies* are sufficient to justify downgrading. Please note that industry funding should not be used as a reason to downgrade a study (section 8.15.1.6 of v5.2 of the Cochrane Handbook[2]).
- Indirectness – judgements can be made with reference to the characteristics of included studies. Please make a clear statement in the main body of the review (Results of the search – Included studies) on whether or not your included studies are sufficient to answer your research question.
- Inconsistency – it is important to consider whether all studies have the same direction of effect, and whether their confidence intervals generally overlap, and not focus solely on the I^2 value. If subgroup analysis has identified possible reasons for the heterogeneity, there is less reason to downgrade than if the heterogeneity remains unexplained. Consider the number of subgroups under investigation and their relevance to the question. It may be appropriate to add a footnote to explain if there is no downgrade.
- Imprecision – Please consider the size of the studies contributing evidence, and whether they are mostly small trials, in addition to the overall number of participants. There is helpful guidance on imprecision and the optimal information size (OIS) in Section 11.2.2 of the Handbook[4]. In addition, section 5.2.4.2 of the GRADE Handbook[5] describes conditions for downgrading and indicates that a sample size of 2000 could be considered adequate even if the OIS is not met. Cochrane Heart does not accept downgrades for imprecision based on TSA methods.
- Reporting bias – usually identified using funnel plots, where appropriate. May also be suspected where a large number of studies in the review do not contribute to an outcome.

---

[3] Sterne JAC, Egger M, Moher D, Boutron I (editors). Chapter 10: Addressing reporting biases. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017), Cochrane, 2017. Available from www.training.cochrane.org/handbook.

[4] Schünemann HJ, Oxman AD, Higgins JPT, Vist GE, Glasziou P, Akl E, Guyatt GH on behalf of the Cochrane GRADEing Methods Group and the Cochrane Statistical Methods Group. Chapter 11: Completing 'Summary of findings' tables and grading the confidence in or quality of the evidence. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). Cochrane, 2017. Available from www.training.cochrane.org/handbook.

[5] http://gdt.guidelinedevelopment.org/app/handbook/handbook.html#h.qoxhi6qajv5t

8. 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 in to the GRADE assessment. Details of funders and any influence they had should be provided in the Characteristics of Included Studies table.

Cochrane Handbook 8.15.1.6 Other issues[2] - Influence of funders

9. Inclusion of 0 event studies in a meta-analysis and forest plot

The problem: If there are 0 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 0 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 0 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.

Nb – this only applies where there are 0 events in both trial arms. For studies with 0 events in one arm but 1 or more events in the other arm, this is not a problem since Revman adds a continuity correction and can include them in the meta-analysis.

10. 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).

11. Imputation of missing data

Please see MECIR standard R44 and the Cochrane Handbook[6] for guidance on handling missing data: https://training.cochrane.org/handbook/chapter-16

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.

12. 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 MI, but you can't assume there weren't any when summarising SAE).

13. Interpretation of results

Many authors find it helpful to use the document on our website "describing results" (https://heart.cochrane.org/for-authors/essential-resources), kindly produced by the Cochrane Consumers and Communication Group.[7]

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 may be preferable to make a distinction between 'no effect' and 'little effect' if the confidence intervals are sufficiently tight around the null point.

14. Statistical significance

It is preferable to avoid describing results as statistically significant (or not) and carefully 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 from Section 12.6.4 of the Handbook (https://training.cochrane.org/handbook/chapter-12)[8]: "**If the confidence interval for the estimate of the difference in the effects of the interventions overlaps the null value, the analysis is compatible with both a true beneficial effect and a true harmful effect. If one of the**

[6] : Higgins JPT, Deeks JJ, Altman DG (editors). Chapter 16: Special topics in statistics. In: Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

[7] 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.

[8] Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, Akl E, Guyatt GH on behalf of the Cochrane Applicability and Recommendations Methods Group. Chapter 12: Interpreting results and drawing conclusions. In: Higgins JPT, Churchill R, Chandler J, Cumpston MS (editors), Cochrane Handbook for Systematic Reviews of Interventions version 5.2.0 (updated June 2017). Cochrane, 2017.

possibilities is mentioned in the conclusion, the other possibility should be mentioned as well."

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.