

Reporting results in CKT reviews

(using material adapted from EPOC and CCCR)

Principles when reporting results

1. Describing and interpreting the results must take into account two key factors:
 - the **certainty** (quality) of the evidence on which the result is based, including the precision of the effect
 - the **size** of the effect (magnitude or importance), including clinical significance
2. The certainty of the evidence should be presented together with effect estimates for each outcome rather than elsewhere in the Results section. Definitions for levels of certainty (high, moderate, low, and very low) can be found in Table 1.
3. Present the results consistently, using similar words and expressions, such as those suggested in Table 2, for similar levels of importance of the effects and certainty of the evidence.
4. Ensure that effects are reported consistently across all sections of the review, including results tables, forest plots, summary of findings tables, the abstract, the plain language summary, the Results section, and the summary of main results in the Discussion section.
5. Include confidence intervals (and P values) when relevant. **Do not report results as being statistically significant or nonsignificant!** (see below)
6. In results tables and the Results section of the review, present results for all the outcomes that are specified in the Methods section.
7. In the abstract, plain language summary, summary of main results in the Discussion section, and summary of findings tables, only present results for the most important outcomes, as specified in the protocol, and try to present no more than seven outcomes.
8. If you found no data for an outcome, present the outcome anyway and note that no data were found.

Certainty of the evidence

It is now **mandatory** that GRADE is used to assess certainty in all new reviews.

Use the term “certainty” (rather than “quality” or “confidence”) throughout your review as it avoids confusion by clearly separating the GRADE assessment from the Risk of Bias assessment.

Table 1: Definitions for level of certainty

GRADE assessment of the certainty of the evidence	Definitions
High	This research provides a very good indication of the likely effect. The likelihood that the effect will be substantially different [†] is low.
Moderate	This research provides a good indication of the likely effect. The likelihood that the effect will be substantially different [†] is moderate.
Low	This research provides some indication of the likely effect. However, the likelihood that it will be substantially different [†] is high.
Very low	This research does not provide a reliable indication of the likely effect. The likelihood that the effect will be substantially different [†] is very high.

[†] *Substantially different = a large enough difference that it might affect a decision.*

Size of the effect

Look at the size of the effect: is it a large or important effect (benefit or harm)? Or a smaller (or negligible) effect?

Note that any judgements made about how important the size of the effect is should be explained in the review.

Standardised statements for reporting effects

One advantage of using GRADE to rate the certainty of the evidence is that it presents the opportunity to use standardised wording, or statements, that reflect the certainty of the evidence.

Table 2 presents standardised wording (statements) which can be used to describe the results and which take into account both the certainty and the importance (size) of the effect.

Because this standardised wording distinguishes between results of greater or lesser quality, and those of more or lesser importance, it gives a matrix of options to allow consistent description of the results across the review.

Selecting the appropriate standardised statement entails three steps, which must be taken for each outcome:

1. Determine the certainty of the evidence of effect for the outcome (Table 1).
2. Determine whether the effect is important, less important, or not important.
3. Go to the corresponding cell in Table 2 and select the appropriate standardised statement.

Table 2: Standardised statements for reporting effects

	Important benefit/harm	Less important benefit/harm	No important benefit/harm
High certainty evidence	[Intervention] improves/reduces/prevents/ leads to [outcome] (high certainty evidence)	[Intervention] slightly improves/slightly reduces/leads to slightly fewer (more) [outcome] (high certainty evidence)	[Intervention] makes little or no difference to [outcome] (high certainty evidence) Or [Intervention] does not have an important effect on [outcome] Or [Intervention] has little or no effect on [outcome]
Moderate certainty evidence	[Intervention] probably improves/reduces/prevents/ leads to [outcome] (moderate certainty evidence)	[Intervention] probably slightly improves/slightly reduces/leads to slightly fewer (more) [outcome] (moderate certainty evidence) Or [Intervention] probably leads to slightly better/worse/less/more [outcome] (moderate certainty evidence)	[Intervention] probably makes little or no difference to [outcome] (moderate certainty evidence)
Low certainty evidence	[Intervention] may improve/reduce/prevent/ lead to [outcome] (low certainty evidence)	[Intervention] may slightly improve/slightly reduce/lead to slightly fewer (more) [outcome] (low certainty evidence)	[Intervention] may make little or no difference to [outcome] (low certainty evidence)
Very low certainty evidence	It is uncertain whether [intervention] improves/reduces/prevents/leads to [outcome] because the certainty of this evidence is very low.		
No data or no studies	[Outcome] was not measured/not reported in the included studies. No studies were found that reported [outcome].		

Examples of statements of effect

- [Intervention A] and [intervention B] may lead to similar health outcomes for [population] (low certainty evidence).
- It is uncertain whether there is any difference between [intervention A] compared to [intervention B] (very low certainty evidence).
- [Intervention A] probably increases [outcome] compared to [intervention B] (moderate certainty evidence).
- It is uncertain whether [intervention] reduces [outcome A] or increases [outcome B] (very low certainty evidence).
- [Intervention] may increase [outcome A] and reduce [outcome B] in [population] (moderate certainty evidence).

Reporting confidence intervals in statements of effects

In situations where the point estimate indicates an important benefit, but the confidence interval is wide and “no effect” or harm is well within the confidence interval, you might want to use the following type of statement:

- **In the abstract and text of the review**

[*Intervention*] may lead to [*better outcome*]. However, the 95% confidence interval indicates that [*intervention*] might make little or no difference / might worsen / increase [*outcome*].

- **In the plain language summary**

[*Intervention*] may lead to [*better outcome*]. However, the range where the actual effect may be (the “margin of error”) indicates that [*intervention*] might make little or no difference / might worsen / increase [*outcome*].

Results should not be reported as statistically significant or statistically non-significant

“Statistical significance” is so commonly misreported and misinterpreted, that we recommend that terms such as ‘not significant’, ‘not statistically significant’, ‘significant’, ‘statistically significant’, ‘trend towards [an effect]’, ‘borderline significant’ should not be used.

For example, authors frequently make a judgement that an intervention works based on the finding of a statistically significant difference between intervention and control groups (e.g. a P value < 0.05 or a confidence interval that excludes no effect). This is misleading as it does not take into account the size of the effect (i.e. is it important?), the precision of the effect estimate, or the quality of the evidence on which it is based.

Similarly, failing to detect a statistically significant result does not necessarily mean that there was no effect. It may be that the result was too imprecise (i.e. too few participants, leading to wide confidence intervals that are consistent with either an important effect or no effect), and measures of statistical significance only give an indication of the likelihood of the result occurring by chance (rather than being due to a real effect). Note that in these cases, the certainty of evidence should have been downgraded for imprecision in the SoF table.

Statistical significance (or lack thereof) should therefore not be used in place of carefully interpreting the size or importance of the effect.

Plain language should be used to describe effects based on the size of the effect and the quality of the evidence. In general point estimates and confidence intervals, when possible, or P-values should be reported, as supporting data for the statements made.

Confusing 'evidence of no effect' with 'no evidence of effect'

When there is inconclusive evidence about the effects of an intervention on an outcome, the conclusion should be that there is 'no evidence of effect'; or that the data suggest that either an increase or decrease in the outcome is possible as a result of the intervention.

It is not correct to conclude that there is 'evidence of no effect' or that an intervention 'showed no effect' because it is possible, for example, that the included studies are too small to detect an effect; or that methodological limitations of included studies mean that an effect has not been detected. Therefore, avoid using 'evidence of no effect' or 'no effect'.

There are other ways that you can describe results, for example:

- 'There is/are currently no evidence/ insufficient data [of / to indicate] an effect of the intervention, compared with control, in terms of effects on [outcome]...'
- 'There is insufficient evidence to decide between intervention and control groups in terms of effects on [outcome]...'
- 'The available evidence is consistent with either an increase or a decrease in [outcome] as a result of the intervention...'

Reporting Risk of Bias

The following wording is recommended for low and unclear risk of bias when completing Risk of Bias tables (see also Appendix 2 in the review document). NOTE: Cells cannot be left empty; this affects the risk of bias figures.

- Random sequence generation
 - Low risk – Provide example from study
 - Unclear risk – 'Study was described as randomised; method of randomisation was not reported'
- Allocation concealment
 - Low risk – Provide example from study
 - Unclear risk – 'Not reported'
- Blinding
 - Low risk – Provide example from study
 - Unclear risk – 'Insufficient information to permit judgement'
- Incomplete outcome data (attrition bias)
 - Low risk – 'No missing outcome data' or 'All patient outcome data reported'
 - Unclear risk – 'Insufficient information to permit judgement'
- Selective reporting
 - Low risk – 'All expected outcomes were reported' or 'Pre-specified outcomes (of interest to this review) were reported'
 - Unclear risk – 'Insufficient information to permit judgement'
- Other bias
 - Low risk – 'The study appears to be free of other sources of bias'
 - Unclear risk – 'Insufficient information to permit judgement'

Reporting funding sources

The funding source is usually listed in two places:

- Characteristics of Included studies table “Notes” section
 - e.g. Funding source: [list the reported funding] or record "not reported"
- Risk of Bias table “Other bias” section
 - Commercial funding = high risk (unless there is an explicit statement saying they had no role in data analysis, interpretation or decision to publish)
 - NGO/not for profit funding = low risk
 - Unspecified = unclear

Common mistakes in presenting or describing the results

Since Cochrane reviews are often large and complex pieces of research, there are many errors that can be introduced when describing and/or interpreting the effects of interventions.

Some of the most common are included in Table 3 below, alongside suggested approaches for good practice, or examples.

Table 3: Common problems in reporting results in Cochrane reviews*

Common problem	Suggested good practice or examples
Inconsistent main messages across sections of the review - particularly the Abstract, SoF tables, PLS, Effects of interventions, Discussion & Implications sections	Use the GRADE ratings as a basis for describing the findings throughout the review
Under-reporting of the primary outcomes and harms, often with emphasis on positive secondary endpoints - particularly within the Abstract	Report the main (primary) outcomes, irrespective of the findings and the strength of evidence. In general, outcomes important enough to have been selected for the SoF tables should be included in the abstract, and vice versa
Little or no use of the information presented in the SoF table in the Abstract, leading to inconsistent messages about effects	Describe the quality of evidence according to GRADE ratings, and ensure consistency with the SoF table(s).

<p>Describing results that are imprecise as being the same as 'no effect' or 'no difference' or 'equally effective'</p>	<p>Highlight the uncertainty in the effects rather than making a judgement about whether the effects are 'present' or 'absent'.</p> <p>For example:</p> <p>'We cannot tell from our results whether the intervention has an important effect on [outcome] because the sample size was small/the results were too imprecise to rule out a small or no effect'</p>
<p>Too much emphasis on statistical significance:</p> <ul style="list-style-type: none"> • A failure to detect a statistically significant effect is misinterpreted as a lack (absence) of an effect • Where a statistically significant result is found, too much emphasis is placed on the presence of an effect 	<p>Emphasise the size (magnitude), the precision (confidence intervals) and the importance of the effect estimate.</p> <p>Integrating the GRADE ratings into the language used to describe results can help to provide a context for the results and to avoid reporting results simply as statistically significant or not (or present and absent).</p>
<p>Wording that associates the quality of evidence with statistical significance</p> <p>For example:</p> <p>'moderate quality evidence of no statistical significance'</p>	<p>Emphasis on the quality of the evidence and the estimate of effect.</p> <p>For example:</p> <p>'The effect of the intervention was uncertain due to imprecision (moderate quality evidence).'</p>
<p>Discussion of the quality of the evidence restricted to considering the risk of bias criteria only, without considering how other factors might impact on quality of evidence (such as imprecision, indirectness, inconsistency and publication bias).</p>	<p>Emphasis on how the GRADE ratings (domains) may influence the findings of key outcome results.</p> <p>Use information about the possible impacts on the quality of evidence than risk of assessments alone.</p>
<p>Very little use of the quality of evidence ratings from SoF tables, information on the decisions about downgrading the evidence, or information about the GRADE methods used.</p>	<p>Refer to, and explain the reasons for downgrading the quality of evidence contained in the GRADE or SoF tables, as needed.</p> <p>Describe the methods used to GRADE the quality of the evidence.</p>

<p>Confusing 'evidence of no effect' with 'no evidence of effect'</p>	<p>Where evidence is inconclusive about the effects of an intervention on an outcome, this represents 'no evidence of effect' (rather than 'evidence of no effect'), i.e. the result suggests that either an increase or decrease in the outcome is possible as a result of the intervention – we are uncertain about the result.</p>
<p>Lack of consistency in the way results are interpreted and reported from one outcome to another</p>	<p>Sometimes similar results are obtained from meta-analysis for different outcomes – e.g. finding uncertain results for both outcome of interest.</p> <p>These results must then be described in consistent ways – e.g. 'we are uncertain about the effects of the intervention on outcome x'; rather than emphasising one finding over another (e.g. stating that effects are uncertain for one outcome, but that there was a small effect for the other outcome); or describing the results with a different emphasis that might no longer be an objective reporting of the findings.</p>
<p>Describing uncertain results as 'no evidence of effect'</p>	<p>Stating 'no evidence of effect' can be misleading as it does not consider the quality of the evidence as an input to deciding how certain we can be about that result, and relies heavily simply on a test of statistical significance. It is preferable to report the result in terms of both the size and quality of the evidence.</p>

<p>Confusing evidence that is poor quality with no evidence</p>	<p>Sometimes evidence that is poor quality can be confused with no evidence, for example:</p> <p>'There is no evidence to decide whether the intervention improves knowledge.'</p> <p>It is really only accurate to state that there is no evidence when no studies were found to measure an outcome; and this statement does not refer to the quality of the evidence (and hence our level of certainty about it).</p> <p>This could be more accurately stated to emphasise that the quality of the evidence is very low and so leads to uncertainty about the effects of the intervention, for example:</p> <p>'As the evidence for our main outcomes is of very low quality, the effects of the intervention on knowledge are uncertain.'</p>
---	--

**adapted from 'Incorporating GRADE in Cochrane reviews: feedback from the CEU screening programme' Lasserson T., Santesso N., Cumpston M., Marshall R., NiÓgáin O. Available at: <http://editorial-unit.cochrane.org/mecir>*

Additional Supporting Material

More information on how to interpret and write results is available at:

- Chapters 11 and 12 of the Cochrane Handbook <http://handbook.cochrane.org/>

More information about interpreting p values is available at:

- Chapter 12 of the Cochrane Handbook <http://handbook.cochrane.org/>
- Effective Practice of Care (EPOC) Author resources. Results should not be reported as statistically significant or statistically non-significant
[http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/22 Interpreting statistical significance 2013 08 12 2.pdf](http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/22%20Interpreting%20statistical%20significance%202013%2008%2012%202.pdf)
- The Cochrane Training website has an interactive training module on interpreting imprecision at <https://training.cochrane.org/resource/interpreting-results-and-drawing-conclusions-online-learning-module> (see section 4 on 'Interpreting results of statistical outputs') and an introduction to meta-analysis and interpreting imprecision at <https://training.cochrane.org/resource/introduction-meta-analysis>
- Wood et al. 2014. Trap of trends to statistical significance: likelihood of near significant P value becoming more significant with extra data. *BMJ* 348: g2215
<http://www.bmj.com/content/348/bmj.g2215>
- McCormack et al. 2013. How confidence intervals become confusion intervals. *BMC Medical Research Methodology*, 13:134 <http://www.biomedcentral.com/1471-2288/13/134>

Using material adapted from

Effective Practice and Organisation of Care (EPOC). Results should not be reported as statistically significant or statistically non-significant. EPOC Resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services; 2013.
[http://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/22%20Interpreting%20statistical%20significance%202013%2008%2012 2.pdf](http://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/22%20Interpreting%20statistical%20significance%202013%2008%2012%202.pdf)

Effective Practice and Organisation of Care (EPOC). Reporting the effects of an intervention in EPOC reviews. EPOC Resources for review authors. Oslo: Norwegian Knowledge Centre for the Health Services; 2016.
[http://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/24. how to report the effects of an intervention 2017 06 01.pdf](http://epoc.cochrane.org/sites/epoc.cochrane.org/files/public/uploads/24.%20how%20to%20report%20the%20effects%20of%20an%20intervention%202017%2006%2001.pdf)

Ryan R, Synnot A, Hill S (2016) Describing results. *Cochrane Consumers and Communication Group.*
http://cccrq.cochrane.org/sites/cccrq.cochrane.org/files/public/uploads/results_wording_revised_december_1st.pdf