Reporting the effects of an intervention in EPOC reviews

Version: 21 June 2018
Cochrane Effective Practice and Organisation of Care Group

Reporting the effects of an intervention in EPOC reviews

Principles when reporting the effects of an intervention

1. In results tables and the Results section of the review, present results for all the outcomes that are specified in the Methods section.
2. 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, and try to present no more than seven outcomes.
3. If you found no data for an outcome, present the outcome anyway and note that no data were found.
4. The certainty of the evidence should be presented together with effect estimates for each outcome rather than elsewhere in the Results section. Use the term "certainty" (rather than "quality" or "confidence") throughout your review. EPOC definitions for levels of certainty (high, moderate, low, and very low) can be found in Appendix 1.
5. Present the results consistently, using similar words and expressions, such as those suggested in Table 1, for similar levels of importance of the effects and certainty of the evidence.
6. Ensure that effects are reported consistently in results tables and forest plots, summary of findings tables, and across the following sections of the review: the abstract, the plain language summary, the Results section, and the summary of main results in the Discussion section.
7. Include confidence intervals or P values when relevant. **Do not report results as being statistically significant or nonsignificant!** (See EPOC resource on interpreting statistical significance)

Using consistent language to report the effects of an intervention

The language that you use to report effects in the text of your review should reflect the importance of the effect and the certainty of the evidence, and it should be consistent throughout the review. The language that is used to describe effects can be difficult to get right. It is easy to cause confusion and misinterpretation by using words inconsistently or by using overly complicated phrases such as "a high likelihood of a somewhat small but possibly important effect". To help authors formulate clear, consistent statements, we present a system of standard statements or wording in Table 1 below. By 'standardized statements' we mean an expression of your results in plain language, using similar words and expressions for similar levels of effect. Table 1 shows which statements to use for different combinations of the importance of the effect and the certainty of the evidence.

If your assessment of the certainty of the evidence is anything other than high, then you should avoid strong statements such as "[intervention] leads to ["outcome"]. You should rather indicate to the reader that there is some degree of uncertainty by adding modifying terms such as "probably", "may" (see Table 1 for suggestions). We acknowledge that the modifying terms we have suggested in Table 1 (such as "probably" and "may") have different meanings to different people and may be difficult to translate into other languages. Nonetheless, the principle of including modifying terms when there is some degree of uncertainty should be adhered to.

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.
2. Determine whether the effect is important, less important, or not important.
3. Go to the corresponding cell in Table 1 and select the appropriate standardised statement.

The EPOC Worksheets for preparing a Summary of Findings (SoF) table using GRADE can be used for making judgements about the certainty of the evidence of effect for each outcome. Guidance for making judgements about whether an effect is important, less important or not important is provided in Appendix 2. These judgments should focus on the importance to people making or affected by a decision. How important an effect is depends on the size of the effect and how much people value the outcome, not on statistical significance.

You may need to amend the statements in Table 1 to fit your intervention or outcome. Any amendments that you make to the statements should be consistent with the suggestions in the table and should be used consistently throughout your review.[1]

Table 1: Standardised statements for reporting effects		
	Important benefit / harm	**Less important benefit / harm**	**No important benefit / harm**
High certainty evidence	[*Intervention*] improves/reduces [*outcome*] (high certainty evidence)	[*Intervention*] slightly improves/reduces [*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 [*outcome*] (moderate certainty evidence)	[*Intervention*] probably slightly improves/reduces [*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 [*outcome*] (low certainty evidence)	[*Intervention*] may slightly improve/reduce [*outcome*] (low certainty evidence)	[*Intervention*] may make little or no difference to [*outcome*] (low certainty evidence
Very low certainty evidence	We don't know if/It is uncertain whether [*intervention*] improves/reduces [*outcome*] because the certainty of this evidence is very low.		
The point estimate	*Option 1* [*Intervention*] may lead to [*better / worse outcome / little or no difference*]. However, the 95% confidence interval		

_____

[1] The development of standard statements or wording to express the results of intervention reviews in plain language is now a GRADE Working Group project, and Table 1 is likely to be updated in response to this work.

indicates an important benefit or harm, and the confidence interval also includes an important benefit / harm / no effect*	indicates that [*intervention*] might make little or no difference / might worsen / increase [*outcome*]. **Option 2** [*Intervention*] may lead to [*better outcome*]. However, the range where the actual effect may be (the "margin of error") indicates that [*intervention*] may make little or no difference / might worsen / increase [*outcome*]. **Option 3** [*Intervention*] may lead to [*better / worse outcome / little or no difference*]. However, the effects of [*intervention*] vary and it is possible that [*intervention*] makes little or no difference / worsens / increases [*outcome*].
No data or no studies	[*Outcome*] was not measured/reported in the included studies. No studies were found that reported [*outcome*]

* See explanation below

Examples of statements of effect (based on summary of findings tables found **here**)
Substitution of nurses for physicians in primary care:
- Care provided by nurses and physicians may lead to similar health outcomes for patients.
- It is uncertain whether there is any difference in the cost of care provided by nurses compared to the cost of care provided by physicians.

Using lay health workers as an add-on to usual care:
- Probably increases immunisation coverage and breast feeding
- May increase care seeking behaviour for children under five and reduce morbidity and mortality in children under five and neonates

Educational meetings for health professionals:
- Probably improve compliance with desired practice and patient outcomes

Introducing user fees for health services in low- and middle-income countries:
- It is uncertain whether introducing user fees reduces health service utilisation or increases inequities in low- and middle-income countries.

Reporting confidence intervals in statements of effects
You will generally not need to refer to the confidence interval or p-value in your statements of effects. This is because your assessment of the certainty of the evidence using GRADE already incorporates a judgement about the precision of effect estimates. However, there may be situations in which it is useful to refer to the confidence interval around an estimate. For instance, in situations where the point estimate indicates a benefit or harm that is likely to be important for decision makers or clinicians, and the confidence interval also includes:

- both an important benefit and an important harm
- "no effect" and an important benefit / harm

In this situation, you might want to use the following type of statement:

Option 1

[*Intervention*] may lead to [*better / worse outcome / little or no difference*]. However, the 95% confidence interval indicates that [*intervention*] might make little or no difference / might worsen / increase [*outcome*].[2]

Option 2

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

Option 3

[*Intervention*] may lead to [*better / worse outcome / little or no difference*]. However, the effects of [*intervention*] vary and it is possible that [*intervention*] makes little or no difference / worsens / increases [*outcome*].

Example 1

Using lay health workers as an add-on to usual care may increase parents' care seeking behaviour for children under five. However, the 95% confidence interval indicates that it may make little or no difference.

Example 2

Using lay health workers as an add-on to usual care may increase parents' care seeking behaviour for children under five. However, the range where the actual effect may be (the "margin of error") indicates that it may make little or no difference.

Example 3

Using lay health workers as an add-on to usual care may increase parents' care seeking behaviour for children under five. However, the effects of using lay health workers vary and it is possible that they make little or no difference to this behaviour.

What are these instructions based on?

This guidance builds on earlier instructions developed by Claire Glenton and Elin Strømme Nilsen (Cochrane Norway), Nancy Santesso (Cochrane Applicability and Recommendations Methods Group) and Simon Goudie and Eamonn Noonan (Campbell Collaboration), and on the following sources:

1. Glenton C, Santesso N, Rosenbaum S, Nilsen ES, Rader, T, Ciapponi A, Dilkes H. Presenting the results of Cochrane systematic reviews to a consumer audience: A qualitative study. Medical Decision Making 2010 Sep-Oct; 30(5):566-77
2. Santesso N, Glenton C, Rosenbaum S, Strømme Nilsen E, Rader T, Pardo J, Ciapponi A, Moja L, Schünemann H. A new format for plain language summaries: does it improve understanding, and is it useful and preferable? A randomised controlled trial. 17th Cochrane Colloquium. Singapore, 2009
3. Glenton C, Kho M, Underland V, Nilsen, ES, Oxman A. Summaries of findings, descriptions of interventions and information about adverse effects would make reviews more informative, Journal of Clinical Epidemiology 2006, 59 (8): 770-778
4. Woloshin S, Schwartz LM. Communicating data about the benefits and harms of treatment: A randomized trial. Annals of Internal Medicine 2011; 155:87-96.

---

[2] This option is not suitable for the plain language summary as the term 'confidence interval' is not understood widely.

Appendix 1: EPOC definitions for levels 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.*

Appendix 2: Deciding whether the size of an effect is important, less important or not important

1. **When judging whether the size of an effect is important, your first step is to establish a threshold - or a range within which the threshold lies (In Figures 1 and 4 you can see examples of this, as suggested by the green shading around the threshold).**

This threshold should focus on the **absolute effect**, since it is difficult, if not impossible to judge the importance of a relative effect alone. For example, a relative risk reduction of 20% (a relative effect) for someone with a 20% likelihood of experiencing a serious (bad) condition would mean a difference (an absolute effect) of 4% or 40 fewer people who would experience that condition per 1000 people (0.20 x 0.20). However, the same relative effect for someone with a 1% likelihood of experiencing the condition would mean a difference of only 0.2% or 2 per 1000 people (0.20 x 0.01), a much less important effect.
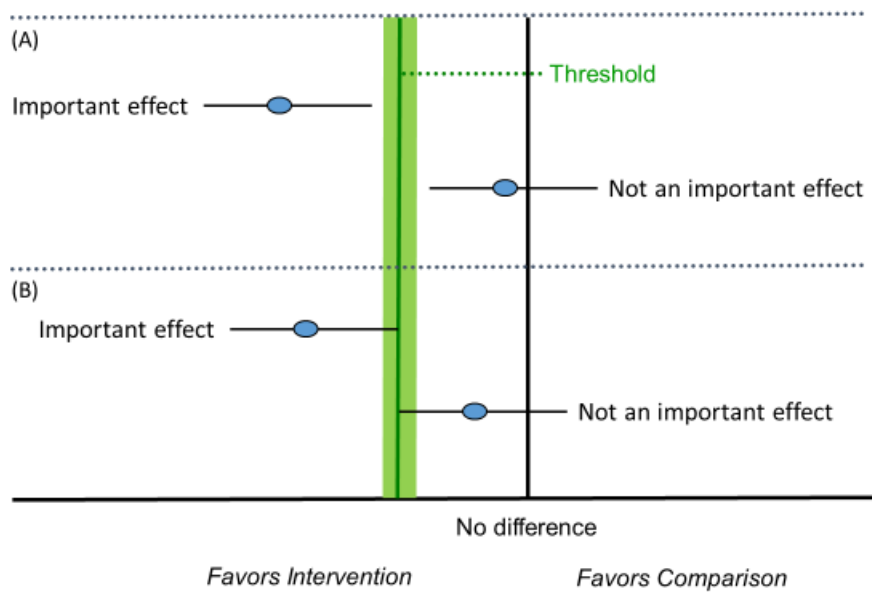
When possible, you should establish in the protocol the threshold for what is considered an important effect, and you should always make it explicit. You should base the threshold on how much the intended beneficiaries of those effects value each relevant outcome and what they would consider to be an important absolute effect. Typically, review authors will not have reliable information on how much people value the main outcomes and there is often wide variation. Consequently, your description of how you made these judgments might be limited to "We did not specify the smallest important difference for outcomes in our protocol for this review. We assessed the importance of effects and the precision of the estimates based on how likely it seemed to us that some people would make different decisions if the true effect was near one end or the other of the 95% confidence interval."

2. **Your next step is to look at the point (best) estimate for the absolute effect and the confidence interval (imprecision of that estimate).**

If the 95% confidence interval does not cross the threshold (or range of values) where you believe that people would go from considering the effect important to considering it unimportant (Figure 1A) or it just touches or barely crosses the threshold (Figure 1B), then the effect should be reported as an important benefit or harm (column 1 in Table 1 above) or as not important, depending on which side of the threshold it is on.

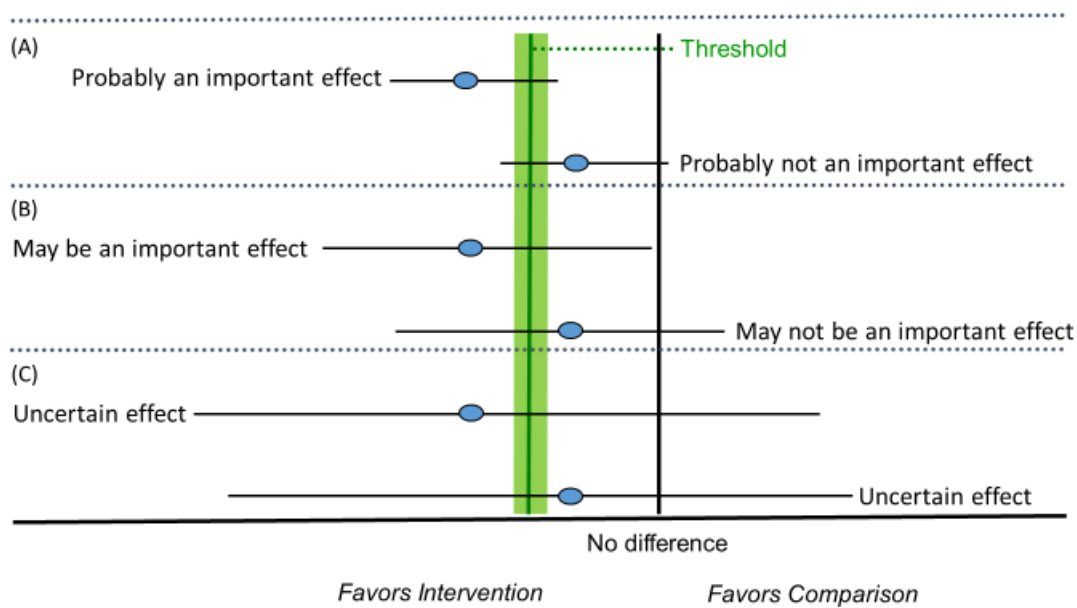**Figure 1. Important effects with little or no serious imprecision**

If the 95% confidence interval crosses the threshold substantially, but is mostly to one side or the other of the threshold, then you should still report the effect as important (or not important). However, you should lower the certainty of the evidence (e.g. from high to moderate (row 2 in Table 1)) if there is serious imprecision (Figure 2A), in which case you should use "probably" when reporting the effect. You should lower the certainty of the evidence from high to low if there is very serious imprecision (Figure 2B), in which case you should use "may" when reporting the effect (row 3 in Table 1).

If there are very little data and very wide confidence intervals (Figure 2C) – i.e., so much imprecision that research does not provide a reliable indication of the likely effect, you should lower the certainty of the evidence from high to very low – then you should simply report the effect as uncertain (row 4 in Table 1).
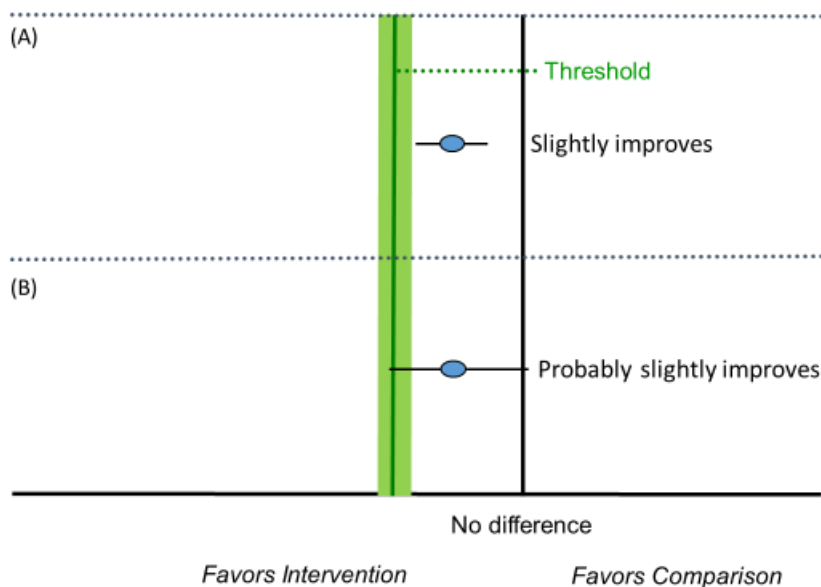
Figure 2. Important effects with serious imprecision

If the 95% confidence interval excludes no difference and excludes the threshold (Figure 3A), then "slightly" should be used when reporting the effect (column 2 in Table 1). If the 95% confidence interval crosses the threshold or no difference, but is mostly to one side or the other of the threshold and no difference, then "slightly" should still be used, but the certainty of the evidence should be lowered; e.g., from high to moderate for serious imprecision, if there are no other reasons for downgrading. In this case "probably" should be used when reporting the effect (Figure 3B).

Figure 3. Less important effects

When the point estimate is close to no difference (column 3 in Table 1), you should **not** report this as "no effect", since it is impossible to rule out any effect. Instead this should be reported as "little or no effect" (or "no important effect"), if there is a narrow confidence interval and no other reasons for downgrading the certainty of the evidence (Figure 4A). If there is a wide 95% confidence interval that does not cross the threshold, then the certainty of the evidence should be lowered; e.g. from high to moderate for serious imprecision, if there are no other reasons for downgrading, in which case "probably" should be used when reporting the effect (Figure 4B).

Figure 4. Little or no effect