

Guidance for better SoF Table explanations and judgments - Cochrane Methods Innovation Fund project

From Santesso et al. “Improving GRADE Evidence Tables part 3: Guidance for useful GRADE certainty in the evidence judgments through explanatory footnotes”, Journal in Clin Epi, in press

Tables

Table 1. Desirable attributes of explanations in Summary of Findings tables (SoF) and Evidence Profiles (EP) – new guidance become available in GRADEpro

Note: guidance and examples refer to explanations only and not if the information is useful for other reasons and in other sections in a review

Attribute and Definition	Key requirements and issues addressed	Examples of acceptable and suboptimal explanations (including examples for differences between SoF tables and EPs)
<p>Concise</p> <p>The explanation should be brief and free from information that is reported in the main text of the review. As a rule of thumb, an explanation should be not longer than two printed lines.</p>	<p>1. Avoid including information that can be included within the table Information about how the outcome was measured or the scale used (in the outcomes column)</p> <p>2. Avoid including information that fits better in the main text Specific details about the analysis or design of included studies. The latter will be covered in the explanation related to risk of bias (if necessary)</p> <p>3. Avoid duplicating information Upon completion of the table, review all explanations to determine if some explanations could be used multiple times if reworded, or combined. Some users may want to keep judgments about the certainty in the evidence separated by domain, in particular in EPs. Since domains may be linked, in specific circumstances more than one judgement about quality can be included in one explanation.</p> <p>4. Avoid redundant information Do not explain that there are no limitations unless the judgment was challenging and users may wonder why it was not addressed.</p>	<p><i>Concise for SoF: Imprecision.</i> The 95% confidence interval includes both no effect and appreciable harm exceeding a minimal important difference.</p> <p><i>Concise for EP: Very serious imprecision.</i> The 95% CI of the pooled relative effect crosses 1 and includes a 3% absolute reduction in death which might be considered important given the low cost and potential harms of the intervention. The optimal information size to detect a minimal important beneficial effect of the magnitude observed here would require > 2,000 patients with 250 events (note: here referring to an outcome associated with fewer events).</p>



<p>Informative</p> <p>The explanation should provide appropriate information to users regarding: 1) judgments; 2) additional sources of information; and 3) interpretation of results</p>	<p>1. Judgements</p> <p>When down- or upgrading is required, authors should provide the rationale for such decision. In addition, they should provide explanations for borderline decisions or justify judgments when there is some concern about the evidence but the certainty rating is not affected. For details about how to provide appropriate explanations for judgements see table 2.</p> <p>2. Additional sources of information</p> <p>This source of information should be related enough to the content of the GRADE table that, in its absence, the autonomy of the table from the main text would be weakened (e.g. where the baseline risk was taken from).</p> <p>3. Knowledge to assist users with interpreting the results</p> <p>This type of explanation should be included when authors anticipate that users would struggle interpreting the results without its inclusion. (e.g. established minimal important difference for a given outcome)</p>	<p><i>Informative: Serious indirectness.</i> One trial included children only, one trial adult soldiers, and one trial adults and children (≥ 14 years). The effect in adults living in the community may be different.</p> <p><i>Not informative: risk of bias.</i> Open-label study; not downgraded for this.</p> <p><i>Informative:</i> Although this was an open label study the outcome assessment appeared free from bias because the investigators protected against bias by blinding outcome assessors and reducing the possibility of altering other care influenced by the open label design.</p> <p><i>Not informative:</i> The treatment difference was not significant.</p> <p><i>Informative:</i> The 95% confidence interval includes both no effect and appreciable harm exceeding a minimal important difference.</p> <p><i>Not informative:</i> RR>2.</p> <p><i>Informative:</i> The quality of evidence was upgraded because of a large effect exceeding a RR of 2 with precise estimates from observational studies that did not suffer from risk of bias or other important limitations.</p> <p><i>Not informative:</i> The control group received placebo.</p> <p><i>Informative:</i> The control group received placebo in 3 out of 5 trials and the effects were similar in the trials using placebo and those that did not. Therefore, we did not downgrade the quality of evidence.</p> <p><i>Not informative:</i> Number needed to treat (NNT) = n/a when result is not statistically significant.</p>
<p>Relevant</p> <p>The explanation should be created keeping in mind the target audience for the GRADE table.</p>	<p>Consider the type of audience your table will inform and how the table will be used for decision-making.</p> <p>In EPs, including the cited references for included studies may be of high importance to the guideline panel while meeting to formulate recommendations. However, if the same references are included in a SoF published accompanying a full systematic review, these can be considered as redundant, unnecessary and irrelevant for users.</p>	<p><i>Relevant:</i> We downgraded two levels because lack of blinding of patients and providers in 4 out of 5 studies; it was unclear if allocation was concealed in 2 studies; and only one study clearly used intention to treat analysis.</p> <p><i>Irrelevant (better placed in methods section of the review):</i> NNT for dichotomous outcomes calculated using NNT calculator (http://www....).</p> <p><i>Irrelevant: No serious inconsistency.</i> The findings of all the trials are consistent. (Note: this information is irrelevant because it is mandatory to consider this criterion and this does not add additional information when the answer to</p>



<p>Easy to understand</p> <p>The explanation should be self-explanatory, not vague, in a way that most users and other target audiences would understand the content provided</p>	<p>After creating your explanation, read it and determine to what extent it is self-explanatory and clear. Consider having a second author check your explanations.</p> <p>If acronyms are used, make sure the full definition is provided</p>	<p>the question about inconsistency is “no”).</p> <p><i>Unclear:</i> Confidence intervals include no difference. <i>Easy to understand:</i> Imprecision is present because the width of confidence interval is consistent with both important benefit and harm.</p> <p><i>Unclear:</i> Although the confidence intervals do not overlap 0.75 and 1.25, the confidence intervals were wide. <i>Easy to understand:</i> Although the confidence intervals do not overlap relative risk estimates of 0.75 and 1.25, the confidence intervals were considered to overlap with the thresholds for decision making. (Note: for EPs, the thresholds should be mentioned)</p> <p><i>Unclear:</i> According to the results of a trial sequential analysis there is firm evidence for a beneficial effect of [drug] versus no placebo or intervention on the [outcome] when the cumulative meta-analysis is adjusted for sparse data and multiple testing on accumulating data. Therefore, there is no risk for random error. <i>Easy to understand for SoF:</i> The confidence intervals are sufficiently narrow for decision-making based on statistical analysis and we did not downgrade. <i>Easy to understand for EP:</i> The confidence intervals are sufficiently narrow for decision-making and we did not downgrade. This is based a trial sequential analysis in which we adjusted for sparse data and multiple testing.</p>
<p>Accurate</p> <p>The content of the explanation should be correct and complete</p>	<p>When the explanation refers to the judgments about the quality of evidence, they should follow the guidance on assessing the certainty in the evidence using the GRADE approach.</p>	<p><i>Inaccurate:</i> Heterogeneity is significantly different. <i>Accurate:</i> There was unexplained inconsistency that was supported by non-overlapping confidence intervals, high I^2 values and statistically significant heterogeneity of effect estimates.</p> <p><i>Inaccurate:</i> The number of trials was too few to assess inconsistency. <i>Accurate:</i> The number of studies was small but no unexplained inconsistency was detected.</p> <p><i>Inaccurate:</i> Downgraded by 1 for imprecision. The presented data appear highly skewed and could not be pooled. <i>Accurate:</i> Certainty in evidence lowered because of a small number of events leading to wide confidence intervals.</p> <p><i>Inaccurate:</i> Downgraded (1 level) because the effect was estimated from a</p>

		<p>single trial. (Note: presence of a single trial is not a reason for imprecision as the trial may be very large, with a sufficient number of events)</p> <p><i>Inaccurate:</i> Wide confidence intervals indicate significant imprecision of this pooled outcome variable, which causes potential bias. (Note: the imprecision causes random error not bias; the statement should have been avoided or refer to imprecision as a reason for downgrading only)</p>
--	--	---

Table 2. Content of explanation for reporting judgments on the certainty in the evidence assessment

GRADE Domain	Relevant content	Example
Criteria for downgrading		
Risk of bias (also known as limitations in detailed study design and execution)	<ul style="list-style-type: none"> - Whether there is serious or very serious issues of risk of bias - Proportion/Number of studies showing shortcomings with any of the risk of bias issues by choosing the appropriate risk of bias tool depending on if randomized or non-randomized studies were use - Contribution of the studies to the pooled estimates 	<p>Risk of bias: Studies that carried large weight for the overall effect estimate rated as high risk of bias due to lack of concealment of randomization and lack of blinding. (For SoF only)</p> <p>Studies that carried large weight for the overall effect estimate rated as high risk of bias due to lack of [concealment] in [3] out of [5] studies and [lack of blinding] in [2] out of [5] studies. (For EP: information in [] can be substituted based on the example)</p>
Imprecision	<ul style="list-style-type: none"> - Whether there is serious or very serious issues of imprecision - Interpretation of the limits of the confidence interval - Determine whether the optimal information size is met 	Serious (or very serious) imprecision. 95% CI is consistent with the possibility for important benefit and large harm exceeding a minimal important difference, including only [] events in total.
Inconsistency	<ul style="list-style-type: none"> - Whether there are serious or very serious issues of inconsistency - Unexplained heterogeneity identified by analysis of point estimates and confidence interval overlap, statistical test (chi-square), and statistical estimates (I^2) 	Serious (or very serious) inconsistency. Unexplained inconsistency, with point estimates widely different and confidence intervals not overlapping (p-value Chi square= []; I^2 =[] %)



Indirectness	<ul style="list-style-type: none"> - Whether or not there are serious or very serious issues of indirectness - Any substantial difference between the identified evidence and the original question of the review regarding: patients, interventions, comparisons, outcomes to the extent that these differences would question the obtained effect estimate - Authors should complete the indirectness table assesses in detail indirectness related to the review or guideline question PICO: population, intervention, comparison and outcome (25, 26) (www.grade-pro.org) 	Serious (or very serious) indirectness. Patients included in the studies have [different condition], and differ importantly from the [question]. Studies also used different [doses of drug] compared to the [question].
Publication bias	<ul style="list-style-type: none"> - Whether publication bias is undetected or suspected - Interpretation of funnel plot - Comprehensiveness of the search strategies and methods to identify all available evidence - Presence of small (often positive) studies with for profit interest 	Publication bias suspected because the included studies were small and the funnel plot shows asymmetry.
Criteria for upgrading		
Large effect	<ul style="list-style-type: none"> - Whether a large effect or association is present and the time frame of exposure necessary to achieve the effect - Explicit description of the magnitude of effect considered as large 	Large effect based on well-done observational studies without important risk of bias or other limitations showing an OR= _____ (95% CI: _____, _____).
Dose-response gradient	<ul style="list-style-type: none"> - Whether the studies provide evidence of a dose-response gradient between intervention or exposure and outcome - Explicit description of the intervention's or exposure's thresholds related to changes (improvement or reduction) in the outcome 	Clear dose-response gradient. RR with the intervention _____ (95% CI: _____, _____) with doses less than _____ and RR of _____ (95% CI: _____, _____) with doses larger than _____.
Opposing plausible residual confounding and bias	<ul style="list-style-type: none"> - Whether the studies provide evidence of all plausible confounders or biases against the detected effect or association, when the later is still detected - Whether the studies provide evidence of all plausible confounders or biases in favour of detecting an effect or association, when the later is not detected - Explicit description of the mechanism for which confounders or biases may be reducing or increasing the observed effect or association 	<p>Confounding and/or biases against the detected effect or association (Description of the mechanism should be provided).</p> <p>Confounding and/or biases in favour of detecting an effect or association not found. (Description of the mechanism should be provided).</p>

Table 3: Guidance for providing useful explanations in Summary of Findings Tables (SoF) and Evidence Profiles (EP)

1. Much of the information you want to communicate to readers can be entered directly into the table and an explanation may not be necessary (e.g. information about the duration of follow-up or the scale used).
2. Generally, do not cite references in the explanations section, unless the GRADE evidence table will be used as a stand-alone EP, for example, in a guideline panel meeting or for providing information about baseline risks.
3. The source of information about the baseline risks used to calculate absolute effects should be provided.
4. Depending on the type of table (SoF, EP, stand alone versus part of a review) explanations may be more or less concise as readers can refer to the review for details.
5. Upon completion of the table, review all of the explanations to determine if some could be referred to multiple times if reworded or combined.
6. Provide reasons for upgrading and downgrading the evidence (see domain specific guidance below).
7. You should consider describing whether you determined that the body of evidence for a particular outcome has serious or very serious issues for the affected domain. Thus, it may be useful to indicate the number of levels for downgrading (e.g. downgraded by one level for risk of bias), but you should avoid repetition of what is in the table and the impression of formulaic or algorithmic reporting. In EPs, this information is already in the cells of the table.
8. Although explanations about the certainty in the evidence are primarily required when they alter the certainty, you could add an explanation to explain when you have not altered the certainty in the evidence, but this decision may be questioned by others. This will help with understanding reasons for disagreement.
9. Remember the table is a summary of the findings not a description of the methods of the review (e.g. do not describe the reasons for the statistical analysis)
10. Results for outcomes that could not be combined statistically in a meta-analysis (i.e. narrative outcomes) can be entered directly into the GRADE evidence table under the results column. An explanation may not be necessary to communicate those results. If you consider that your audience would benefit from adding complementary estimates of treatment effect (e.g. number needed to treat for benefit and harm, risk difference expressed as percentage, continuous outcome expressed in minimal important difference units), these can be included in the comment column.
11. Use the information presented in the explanations in SoF about the GRADE process to develop other key parts of the systematic review, including summary versions and discussion section (24).

Domain specific guidance for writing useful explanations

Risk of bias

1. Qualify the number of studies, or the amount of information that they provide in the meta-analysis, which were at high risk of bias and indicate for which criterion.
 - a. Use terms such as majority, minority, all, some, or none; or indicate the number as X/X studies.
 - b. For randomized controlled trials refer to the specific criteria including allocation concealment, selective outcome reporting, etc. For non-randomized studies refer to the criterion in the tool used (e.g. using the ACROBAT-NRS tool). In some instances, also indicate the contribution of the studies to the estimates.
 - c. Indicate if the effect of the risk of bias was tested in a sensitivity analysis.
2. Information about study design may be included in the explanations, in particular, in SoF when different study designs are included. However, this information is included in the EP.



Imprecision

1. Indicate whether the sample size or number of events does not meet the optimal information size as calculated, or the ‘rules of thumb’ (e.g. >400 events). Avoid reference to the number of studies as a reason for imprecision.
2. Indicate whether the confidence intervals include the possibility of a small or no effect AND important benefit or harm. If the threshold of important benefit is known, provide the numerical value.
3. Avoid reporting the result as statistically or non-statistically significant.

Inconsistency

1. Indicate if the judgement of inconsistency is based on statistical tests (I^2 , Chi^2 , Tau), overlap of confidence intervals, or similarity of point estimates.
2. If inconsistency is based on I^2 , consider describing it as considerable, substantial, moderate or not important.
3. If applicable, describe whether heterogeneity was explored in subgroup analyses by PICO (patients, intervention, comparison, outcome), and indicate other potential reasons for the heterogeneity.
4. In the case of a single study for an outcome, do not indicate that inconsistency is ‘not applicable’. Indicate that there is none.

Indirectness

1. Describe if indirectness is due to the elements of PICO (patients, intervention, comparison, outcome).

Publication bias

1. Indicate the reason publication bias is detected (e.g. asymmetrical funnel plot, small studies with positive results, suspected selective availability of data from published or unpublished studies).

Upgrading

1. In SoF compared to EP, the reasons for upgrading are not provided, therefore write the specific reason: due to large effect; a dose response gradient; or plausible confounding increases confidence.
2. In both SoF and EP indicate the specific reason. For large effects report if relative effect is >2 or >5. For dose response gradients, describe the level of intervention and effect on outcome. For confounding, describe the effect of the confounding factor on the estimate.