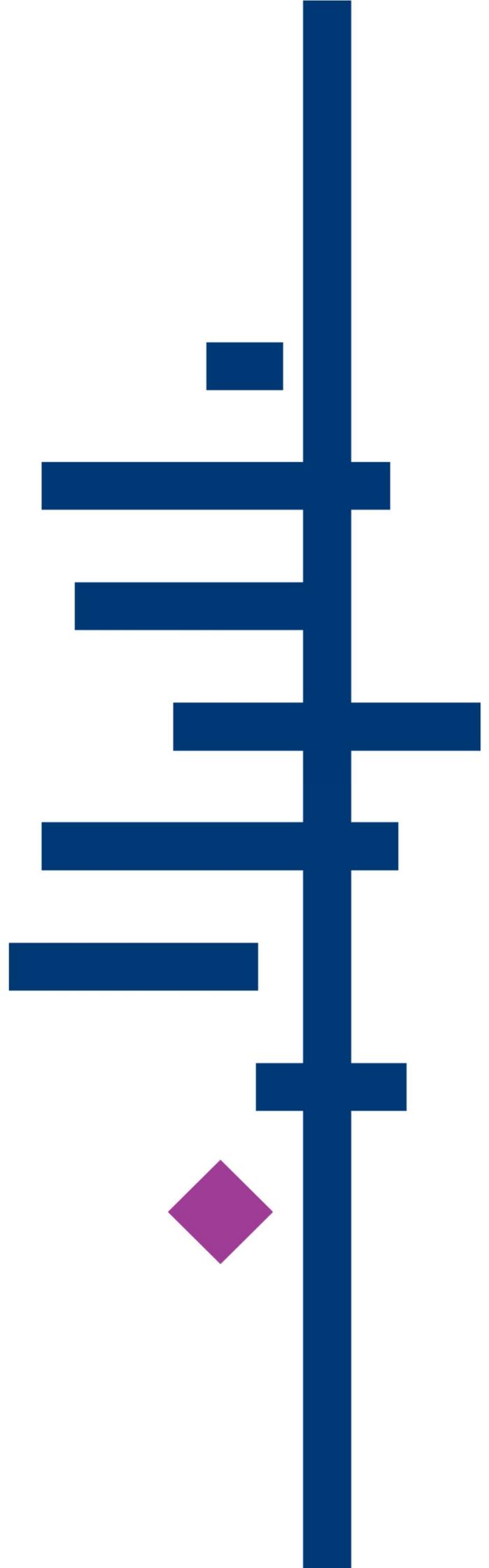


Operational guide to the Cochrane Editorial & Methods Department Review Triaging Tool

Date: 28th February 2019



Contents

1	Introduction	3
2	Triage Tool	3
2.1	Implementation of protocol methods	4
2.2	Summary of Findings Table	7
2.3	Abstract	9
3	Development & Maintenance of the Tool	12
4	Appendix: EMD Pre-publication Review Triage Tool	13

1 Introduction

This operational guide accompanies the Editorial & Methods Department (EMD) pre-publication triage tool for evaluating systematic reviews under consideration for editorial approval in Cochrane. The questions in the tool are designed to elicit yes/no responses, but further comments are encouraged for 'No' decisions and for marginal calls.

The EMD uses this tool on reviews that have been sent for screening. It is primarily intended to help decide what sort of additional work might be needed before publication. It should not be used in isolation to identify major problems. By making the criteria and their rationale transparent we would like to enable others to carry out these checks earlier in CRG editorial processes where feasible.

The tool focuses on three separate aspects of the review: how protocol methods were implemented, the use of GRADE and Summary of Findings (SoF) tables, and the consistency of reporting in the abstract. It adapts the tool that was used in the [2016 abstract audit](#). Whilst there is no separate item relating to readability, clarity of writing should feature as part of the feedback if this is not otherwise covered by specific aspects of reporting.

2 Triage Tool

The tool can be found in the [Appendix](#) of this guide. It is not necessary to use it according to the order of the items that follow. Users are strongly encouraged to work through the tool in the order which they feel best fits their own method of assessment. Our experience to date has led us to assess the first three abstract items before any others to understand the review question and currency of the findings. We then focus on the implementation of protocol methods and analysis items as the foundations on which the interpretation and summary versions of the review are based.

When using the tool make a note of the version of the review being assessed as given in Archie. This will help to track subsequent changes to the review on Archie.

2.1 Implementation of protocol methods

This section of the tool is intended to help users look for features that indicate potential problems with the way that the review protocol was implemented. The tool is oriented around the following items:

- eligibility decisions & risk of bias judgments,
- matching analysis settings & stated methods,
- accounting for study design & double-counting,
- possible transcription errors.

The tool should not be used on its own to diagnose mistakes in the analysis of data. In most cases it will be important to follow up on issues in the analysis by looking at the primary trial reports. Consultation with someone with appropriate statistical expertise might also be necessary.

Start by using Archie to generate a comparison document to identify changes between the published protocol and the review version. This will help judge the alignment of the protocol methods with what was subsequently implemented.

- **Appropriate eligibility decisions**

Pay close attention to any changes to the Population, Interventions/Comparators, or the definition and priority of the Outcomes. Any changes should be clearly justified in the section 'Differences between protocol and review'. Check the 'Characteristics of excluded studies table' and the 'Description of excluded studies' in the text. The reasons for exclusion should align with the review eligibility criteria. Excluding studies because outcomes were not reported may introduce reporting bias in to the review if there is evidence that the outcomes were measured by the studies. Encourage the authors to be more explicit if "outcomes not reported" is given as a justification for exclusion. For example, were relevant outcomes not measured because the objective of the study differed to that of the review objectives? Also check that sufficient explanation and detail is given for any studies awaiting assessment or any ongoing studies.

- **Appropriate risk of bias assessment**

Check that all the standard domains are considered in the risk of bias tables. Be aware that in cases where blinding is not possible the domains relating to blinding should not simply be ignored. Check that the domains appear well understood: the explanation should be appropriate to the domain under

consideration and the judgement of risk of bias should be in line with the explanation. The inclusion of any non-standard domains should be explained and justified. For particularly large reviews it might be best to focus on the tables of the first five or six studies listed. If there are any problems with specific domains (e.g. confusion between concealment of allocation and blinding, or between missing data and selective outcome reporting), then consider checking the other risk of bias tables in more detail.

- **Analyses match with methods section**

Start with the analyses that are used for any of the SoF table outcomes. Compare the default settings in the analysis tables against ‘Measures of treatment effect’ and ‘Data synthesis’ sections in the methods. Check for declared ‘Differences between protocol and review’ to see if any relevant changes to the protocol are explained. Anticipating the nature of the data at the outset of a systematic review can be difficult, so discrepancies between analysis settings and planned methods might simply reflect reasonable but undeclared changes to the protocol methods. Clarification should be sought from the authors in these circumstances.

Where subgroup analyses have been carried out, they should be consistent with those listed under the methods. Check that tests for interaction support assertions about subgroup effects and are interpreted cautiously. If the subgroup function has been used to display different outcomes or timepoints and there is no intention to combine data from them, the totals and test for subgroup differences need to be switched off. Decisions not to implement planned subgroup analyses are reasonable if there are only a few studies with available data, or the classification of studies by subgroup has not been possible.

- **Data from non-standard designs (cluster, crossover, etc.) appropriately incorporated where relevant**
- **Multiple measurements from studies with more than one eligible comparator handled appropriately**

The subsection ‘Unit of analysis issues’ should describe methods for handling studies with a cluster, crossover or other designs where body parts are randomised. It should also include information about handling data from studies with more than one eligible treatment arm. Check this section against what was described in the protocol. If the review question allows for the inclusion of studies with a non-standard design or multiple treatment arms, read the description of included studies in the text and scan the characteristics of included studies to identify any studies where a unit of analysis issue might arise.

For multi-arm studies, this will indicate how many treatment arms are eligible. Look at the data and analysis tables to see whether adjusted data were collected or necessary adjustments were made manually as outlined in 'Unit of analysis issues'. Study arms can be combined or split by the number of relevant comparisons. It is worth looking to see if footnotes have been used in the forest plots to explain data sources used and the nature of any adjustments made.

- **Outlying results acknowledged and explored appropriately**
- **Assessment for data entry errors**

Consider the plausibility of the results overall, but especially where studies have large or discordant effects. Check their influence in the analyses by temporarily deselecting them or sorting the studies by effect or weight in RevMan. Even if the review authors have explained and checked unusual effects, it will be helpful to follow-up by examining what was published in the trial report whenever possible.

Studies reporting continuous data from the same scale should have similar standard deviations (SDs), any SDs that look very small for the scale of interest. There will always be some degree of variation across studies, especially where populations recruited to the studies vary due to severity of condition, study design or setting. Differences between study SDs might also occur where mean differences (MDs) combine change and end of treatment scores. Note that the Handbook advises against combining end of treatment and change from baseline as standardised mean differences (SMDs).

Errors introduced by transcription or transformation can explain why some studies have very small SDs relative to others in the analysis. A standard error might have been entered instead of a standard deviation, but if data have been transformed to a common scale (for example hours to minutes), SDs might have been taken from the original scale and simply not converted.

Errors relating to SDs will show up differently in the analysis depending on the effect measure. For MDs, the weight of the study in the analysis will be greater than you would expect given the sample size. For SMDs the effect size itself will be exaggerated because the SDs used as the basis for estimating the difference in treatment group means will be erroneously small.

The sample size for a study should be similar across different outcomes it contributes to in a review, although due to attrition, sample sizes can get smaller as later timepoints are analysed. Be alert when

the number of participants exceeds the number recruited to a study. This could mean that instead of using the number of people in the treatment groups, the authors have entered another denominator. This might be the number of body parts or even the number of episodes (colds, headaches, or exacerbations) rather than the number of participants. Continuous data analysis methods in RevMan assume that the sample sizes reflect the number of participants.

Any changes to the data in the analyses could impact on the overall results and interpretation. Correcting statistical errors can change the weight of studies at high or low risk of bias, or the size, precision or consistency of effect. It is always worth reconsidering the GRADE ratings following any changes to the data.

2.2 Summary of Findings Table

SoF tables convey key details about the review and the findings. They should tell you about the review question, setting of the studies, results and certainty of evidence for important outcomes, including information on downgrading decisions.

- **SoF table presents main outcomes (both benefits and adverse effects) for main comparison**

SoF tables should report on results for outcomes that are most useful for decision makers. Verify that the list of outcomes presented in the SoF table matches the list of SoF table outcomes in the methods, regardless of the availability of data for each outcome. Both benefits and harms need to be included in the table. Empty rows can be used to represent outcomes where no results are available.

- **PICO (including Settings) presented and accurate**
- **Outcomes fully defined (i.e. time of measurement, scale of measurement, range of scores specified)**
- **Assumed and Corresponding risks presented (where appropriate)**

The SoF table should present a clear and succinct outline of the review question. It should carry information about the participants, comparator interventions and setting as per GRADE guidance.

A common error in SoF tables is the omission of detail about the measurement of outcomes. Information about scales for continuous data are useful because this information is not always obvious from the

results. The duration of follow-up is often overlooked so check that this has been included. Check to see if control group means are presented in the tables. This often needs to be added manually, especially when data have been analysed with generic inverse variance (see [SoF table screening note](#) for further guidance).

- **GRADE ratings justified & adequately explained**

Have the forest plots to hand when checking the GRADE ratings. GRADE ratings should be reasonable with appropriate justifications provided by footnotes. If sensitivity analyses have not been reported by risk of bias, include the risk of bias judgments in the forest plots to see how much influence studies at high risk of bias have on the results. Consider this alongside any downgrading decisions relating to risk of bias.

The number of downgrades should match with the rating (e.g. moderate) and symbols ($\oplus\oplus\oplus\ominus$). The footnotes should convey why the evidence was downgraded and by how many levels, with a summary explanation for the decision. For example: “*Downgraded one level due to imprecision (small number of events and wide confidence interval).*” More detailed guidance on explaining downgrading decisions [is available](#). Upgrading evidence which has already been downgraded is not in line with GRADE guidance. This applies to evidence from both randomised and non-randomised studies.

A general observation is that downgrading decisions should not just be left in the SoF tables. They should be included in other parts of the review. The discussion subsection ‘Quality of the evidence’ or ‘Certainty of the evidence’ sometimes just repeats information about the risk of bias of the studies from the results section. This section should be viewed as an opportunity to summarise the downgrading decisions presented in the SoF tables.

- **Clear and accurate summary of narrative results (where appropriate)**
- **Quality ratings presented for narrative results (where appropriate)**

SoF tables should convey results for the most important outcomes irrespective of the amount of information. It might be that for some outcomes no meta-analysis has been possible, but data from one or more studies have been summarised narratively elsewhere in the review. If such outcomes are

important enough to include in the SoF table, a brief descriptive summary of their results will still be informative (see [SoF table screening note](#) for further guidance). Including this information here means that there is something to work back in to the abstract and plain language summary.

2.3 Abstract

The items in the abstract domain of the tool focus on the research question, search date, results for benefits and harms and main conclusions. The abstract is a standalone summary and point of entry in to the review for the reader. The plain language summary should also address these considerations.

- **Title reflects the review question**
- **Research question (PICO) is clear and the rationale for the review is well described**
- **Search date is less than 12 months from publication**

The title should link to information presented in the background, objectives and eligibility criteria of the abstract. Check that the background succinctly describes the condition or problem addressed by the review question, and provides context for the review. The criteria for selecting studies in to the review should follow logically from the Background and Objectives.

The search date is usually reported in the abstract search methods. Verify this date against the search date field of the review. If the search has identified ongoing studies, scan their characteristics to see if large studies are likely to have completed and even published recently. It is possible to check on publication status by entering study register IDs in to PubMed.

- **Characteristics of included studies summarised**
- **Findings for all important outcomes reported for the main comparison(s), including information about harm (i.e. consistent with the outcomes reported in the SoF table)**

The abstract is an opportunity to draw attention to noteworthy features of the included studies in terms of where they were done, who they recruited and what interventions were used. Summarising these characteristics of the included studies will help readers gauge the applicability of the review findings. Restrictions on word count mean that the detail provided elsewhere in the review needs to be summarised and not copied verbatim. If this information is missing, look for key statements about the

studies under Results>Description of Included Studies, SoF tables, or Discussion>Overall completeness and applicability. These can be copied up to the abstract and edited back as appropriate.

The abstract should closely mirror information in the SoF tables. Both sections should tell you what the findings are for the most important outcomes for the comparison(s) that most directly address(es) the review objectives. When considering completeness and consistency of reporting in the abstract, start with the first SoF table because it will display under the abstract on publication. If the review has many SoF tables there should be some sense of priority to the comparisons.

The reporting of outcomes in the abstract should match with what is listed in the SoF tables. One of the most common errors in abstracts is the omission of adverse events and other important outcomes from the SoF table. This can be because there are either limited or no data or narratively summarised results. Abstracts need to provide a complete summary of results for key outcomes and acknowledge gaps where they exist.

- **Direction, magnitude and confidence intervals of effects clearly described where appropriate**
- **Reporting of results avoids emphasizing statistical significance to determine presence or absence of an effect**
- **GRADE certainty (or quality) of evidence reported for outcomes in the abstract**
- **Absolute effects used to illustrate the relative effects where appropriate**

From reading the abstract it should be possible to understand whether the effect of the intervention was beneficial, harmful or indeterminate for the outcomes of interest. For outcomes measured with continuous data the units and direction of benefit are not always obvious so further explanation might be needed. An increasing number of abstracts use the plain language statements (incorporating both the importance (size) of effect and quality of evidence) to communicate key results (see: [Cochrane Norway website](#)). This is a useful way to place greater emphasis on the direction and size of effect, and to reduce reliance on results being described in terms of statistical significance.

Look for outcomes where it would be helpful to illustrate relative effects with absolute effects. Absolute effects can be sourced from the SoF table columns for control and intervention groups. Note that it may not always be useful to include this information where there is a high degree of uncertainty over the

intervention effect (i.e. very low-quality evidence). It is better to emphasize lack of data or certainty than to place this kind of emphasis on results with limited applicability.

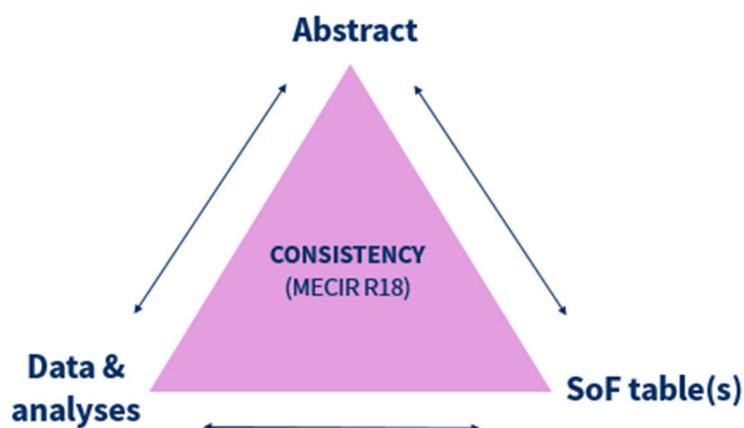
- **Conclusions accurately reflect evidence presented in SoF table(s) & avoid recommendations**

The conclusions should follow from the information presented in the SoF tables. Be alert to how uncertainty is conveyed in the summary versions. Watch out for use of the term ‘safe’ in the review abstract, implications for practice and plain language summaries to describe the adverse event profile of the intervention of interest. This is a term that conveys a particularly reassuring message yet it is rarely justified by the nature of evidence from most intervention reviews.

MECIR guidance is clear that recommendations to adopt or avoid an intervention are to be avoided, so look for coercive language such as ‘should be used/implemented/offered’. By implication, language such as ‘cannot be recommended’ suggests that a recommendation would have been made based on better evidence and this will require rewording.

- **Key findings consistent across the summary versions of the review**

Conclusions should be consistent between the different sections of the review. Use the split text view in RevMan to cross-check the conclusions in the abstract, PLS and full text with information in the SoF tables. Check that numerical results reported in the SoF tables are consistent across the Data and Analysis section, Abstract and PLS. Try to think of the links between the SoF table, analyses and results and abstract as illustrated below:



3 Development & Maintenance of the Tool

The tool has been developed drawing on the experience from the EMD pre-publication screening of Cochrane intervention reviews since 2013. It will undergo periodic update to keep up to date with new developments in the standards for the conduct and reporting of Cochrane reviews.

The tool and this guide have been developed by the Associate Editors in the EMD (Liz Bickerdike, Sarah Hodgkinson, Nuala Livingstone & Newton Opiyo), with input from Toby Lasserson (Senior Editor) and Kerry Dwan (Statistical Editor).

4 Appendix: EMD Pre-publication Review Triage Tool

Review title					
Authors					
CRG					
Archie version no.					
IMPLEMENTATION OF PROTOCOL METHODS		SUMMARY OF FINDINGS TABLE		ABSTRACT	
Item	Response	Item	Response	Item	Response
Appropriate eligibility decisions <i>Check protocol comparison generated from Archie and Differences between protocol & review for any changes to design of review (eligibility criteria, outcomes); Check for exclusions based on reporting of data</i>		SoF table presents main outcomes (benefits & harms) for main comparison <i>Look at methods section for consistency of SoF table outcomes; Assess methods for using GRADE</i>		Title reflects review question Research question (PICO) clear & rationale for review described	

<p>Appropriate risk of bias assessment</p> <p><i>Check for omission of standard domains; inclusion of any non-standard domains is explained & justified; domains appear well understood (fit between explanation and domain, appropriate judgments)</i></p>		<p>PICO (including Settings) are accurate & informative</p>		<p>Search date <12 months from publication</p>	
		<p>Outcomes fully defined (i.e. time of measurement, scale of measurement, range of scores specified)</p>		<p>Characteristics of included studies summarised</p> <p><i>Consider copying & summarising information presented under 'Description of included studies'/Overall completeness & applicability'. Look for details in SoF table relating to settings & participants.</i></p>	
<p>Analyses match with methods section</p> <p><i>MDs/SMDs; fixed/random effects, subgroup analysis. Check protocol comparison and Differences between protocol & review to see what plans changed from protocol.</i></p>		<p>Assumed & Corresponding risks included (where appropriate)</p>		<p>Findings for all important outcomes reported for main comparison(s), including information about harms</p> <p><i>Check consistency with first SoF table & others as appropriate</i></p>	
<p>Data from non-standard designs (cluster, cross-over, etc.) appropriately incorporated where relevant</p> <p><i>Check 'Unit of analysis issues' in methods/footnotes in forest plots/sensitivity analyses. Scan study characteristics to confirm</i></p>		<p>GRADE ratings justified & adequately explained</p>		<p>Direction, magnitude & confidence intervals of effects clearly described where appropriate</p> <p><i>Check units for continuous outcome data</i></p>	

<i>unit of allocation & sample sizes if in doubt.</i>					
<p>Multiple measurements from studies with more than one eligible comparator handled appropriately</p> <p><i>Check for double counting of studies in Forest plot & adjustment of events/sample size in control groups</i></p>		Clear & accurate summary of narrative results (where appropriate)		Reporting results avoids emphasizing statistical significance to determine presence or absence of an effect	
<p>Outlying results acknowledged & explored appropriately</p> <p><i>Consider how plausible the direction/size of effects are overall, explore data from studies with unusually large or discordant effects</i></p>		Quality ratings presented for narrative results (where appropriate)		GRADE ratings for outcomes reported in abstract	
				Absolute effects used to illustrate the relative effects where appropriate	
<p>Assess the following for possible data entry errors:</p> <p><i>Comparability of SDs for studies using same scale</i></p> <p><i>Weights of individual studies</i></p>				Conclusions accurately reflect evidence presented in SoF table(s) & avoid recommendations	

<i>look relative to sample size (high/low SDs or counts)</i>				Key findings consistent across the summary versions of the review	
<i>Similarity of sample sizes across different outcomes for same study</i>					