

Checklist for making judgements about how much confidence to place in a systematic review of effects (adapted version of SURE checklist:)ⁱ

Assessed by:
Date:

Section A: Methods used to identify, include and critically appraise studies

<p>A.1 Were the criteria used for deciding which studies to include in the review reported?</p> <p>Did the authors specify:</p> <p><input type="checkbox"/> Types of studies</p> <p><input type="checkbox"/> Participants/ settings/ population</p> <p><input type="checkbox"/> Intervention(s)</p> <p><input type="checkbox"/> Outcome(s)</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><i>Coding guide - check the answers above</i> <i>YES: All four should be yes</i> <i>NO: All four should be no</i> <i>PARTIALLY: Any other</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>A.2 Was the search for evidence reasonably comprehensive?</p> <p>Were the following done:</p> <p><input type="checkbox"/> Language bias avoided (no restriction of inclusion based on language)</p> <p><input type="checkbox"/> No restriction of inclusion based on publication status</p> <p><input type="checkbox"/> Relevant databases searched (<u>Minimum criteria:</u> All reviews should search at least one source of grey literature such as Google; for health: Medline/ Pubmed + Cochrane Library; for social sciences IDEAS + at least one database of general social science literature and one subject specific database)</p> <p><input type="checkbox"/> Reference lists in included articles checked</p> <p><input type="checkbox"/> Authors/experts contacted</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Can't tell</p> <p><i>Coding guide - check the answers above:</i> <i>YES: All five should be yes</i> <i>PARTIALLY: Relevant databases and reference lists are both reported</i> <i>NO: Any other</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>A.3 Does the review cover an appropriate time period?</p> <p><i>Is the search period comprehensive enough that relevant literature is unlikely to be omitted?</i></p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> Can't tell (only use if no information about time period for search) <input type="checkbox"/> No <input type="checkbox"/> Unsure</p> <p><i>Coding guide:</i> YES: Generally this means searching the literature at least back to 1990 NO: Generally if the search does not go back to 1990 CAN'T TELL: No information about time period for search</p> <p><i>Note: With reference to the above – there may be important reasons for adopting different dates for the search, e.g. depending on the intervention. If you think there are limitations with the timeframe adopted for the search which have not been noted and justified by the authors, you should code this item as a NO and specify your reason for doing so in the comment box below. Older reviews should not be downgraded, but the fact that the search was conducted some time ago should be noted in the quality assessment. Always report the time period for the search in the comment box.</i></p>
<p><i>Comments (note search period, any justification provided for the search period, or uncertainty)</i></p>	

<p>A.4 Was bias in the selection of articles avoided?</p> <p>Did the authors specify:</p> <p><input type="checkbox"/> Independent screening of full text by at least 2 reviewers</p> <p><input type="checkbox"/> List of included studies provided</p> <p><input type="checkbox"/> List of excluded studies provided</p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No</p> <p><i>Coding guide:</i> YES: All three should be yes, although reviews published in journals are unlikely to have a list of excluded studies (due to limits on word count) and the review should not be penalised for this. PARTIALLY: Independent screening and list of included studies provided are both reported NO: All other. If list of included studies provided, but the authors do not report whether or not the screening has been done by 2 reviewers review is downgraded to NO.</p>
<p><i>Comments (note important limitations or uncertainty):</i></p>	

<p>A.5 Did the authors use appropriate criteria to assess the quality and risk of bias in analysing the studies that are included?ⁱⁱ</p> <p><input type="checkbox"/> The criteria used for assessing the quality/ risk of bias were reported</p> <p><input type="checkbox"/> A table or summary of the assessment of each included study for each criterion was reported</p> <p><input type="checkbox"/> Sensible criteria were used that focus on the quality/ risk of bias (and not other qualities of the studies, such as precision or applicability/external validity). “Sensible” is defined as a recognised quality appraisal tool/ checklist, or similar tool which assesses bias in included studies. Please see footnotes for details of the main types of bias such a tool should assess.</p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No</p> <p><i>Coding guide:</i> <i>YES: All three should be yes</i> <i>PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY.</i> <i>NO: Any other</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

Section B: Methods used to analyse the findings

<p>B.1 Were the characteristics and results of the included studies reliably reported?</p> <p>Was there:</p> <p><input type="checkbox"/> Independent data extraction by at least 2 reviewers</p> <p><input type="checkbox"/> A table or summary of the characteristics of the participants, interventions and outcomes for the included studies</p> <p><input type="checkbox"/> A table or summary of the results of all the included studies</p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially <input type="checkbox"/> Not applicable (e.g. no included studies)</p> <p><i>Coding guide:</i> <i>YES: All three should be yes</i> <i>PARTIALLY: Criteria one and three are yes, but some information is lacking on second criteria.</i> <i>No: None of these are reported. If the review does not report whether data was independently extracted by 2 reviewers (possibly a reporting error), we downgrade to NO.</i> <i>NOT APPLICABLE: if no studies/no data</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>B.2 Are the methods used by the review authors to analyse the findings of the included studies clear, including methods for calculating effect sizes if applicable?</p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data)</p> <p><i>Coding guide:</i> <i>YES: Methods used clearly reported. If it is clear that the authors use narrative synthesis, they don't need to say this explicitly.</i> <i>PARTIALLY: Some reporting on methods but lack of clarity</i> <i>NO: Nothing reported on methods</i> <i>NOT APPLICABLE: if no studies/no data</i></p>

<i>Comments (note important limitations or uncertainty)</i>	
<p>B.3 Did the review describe the extent of heterogeneity?</p> <p><input type="checkbox"/> Did the review ensure that included studies were similar enough that it made sense to combine them, sensibly divide the included studies into homogeneous groups, or sensibly conclude that it did not make sense to combine or group the included studies?</p> <p><input type="checkbox"/> Did the review discuss the extent to which there were important differences in the results of the included studies?</p> <p><input type="checkbox"/> If a meta-analysis was done, was the I^2, chi square test for heterogeneity or other appropriate statistic reported? If no statistical test was reported, is a qualitative justification made for the use of random effects?</p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data)</p> <p><i>Coding guide:</i> YES: First two should be yes, and third category should be yes if applicable should be yes PARTIALLY: The first category is yes NO: Any other NOT APPLICABLE: if no studies/no data</p>
<i>Comments (note important limitations or uncertainty)</i>	

<p>B.4 Were the findings of the relevant studies combined (or not combined) appropriately relative to the primary question the review addresses and the available data?</p> <p>How was the data analysis done?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Descriptive only <input type="checkbox"/> Vote counting based on direction of effect <input type="checkbox"/> Vote counting based on statistical significance <input type="checkbox"/> Description of range of effect sizes <input type="checkbox"/> Meta-analysis <input type="checkbox"/> Meta-regression <input type="checkbox"/> Other: specify <input type="checkbox"/> Not applicable (e.g. no studies or no data) <p>How were the studies weighted in the analysis?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Equal weights (this is what is done when vote counting is used) <input type="checkbox"/> By quality or study design (this is rarely done) <input type="checkbox"/> Inverse variance (this is what is typically done in a meta-analysis) <input type="checkbox"/> Number of participants (sample size) <input type="checkbox"/> Other: specify <input type="checkbox"/> Not clear <input type="checkbox"/> Not applicable (e.g. no studies or no data) <p>Did the review address unit of analysis errors?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Yes - took clustering into account in the analysis (e.g. used intra-cluster correlation coefficient) <input type="checkbox"/> No, but acknowledged problem of unit of analysis errors <input type="checkbox"/> No mention of issue <input type="checkbox"/> Not applicable - no clustered trials or studies included 	<p><input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data) <input type="checkbox"/> Can't tell</p> <p><i>Coding guide:</i> YES: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors addressed (if appropriate). PARTIALLY: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors not addressed (and should have been). NO: If narrative OR vote counting (where quantitative analyses would have been possible) OR inappropriate reporting of table, graph or meta-analyses. NOT APPLICABLE: if no studies/no data CAN'T TELL: if unsure (note reasons in comments below)</p>
---	--

<i>Comments (note important limitations or uncertainty)</i>	
<p>B. 5 Does the review report evidence appropriately?</p> <p><input type="checkbox"/> The review makes clear which evidence is subject to low risk of bias in assessing causality (attribution of outcomes to intervention), and which is likely to be biased, and does so appropriately</p> <p><input type="checkbox"/> Where studies of differing risk of bias are included, results are reported and analysed separately by risk of bias status</p>	<p><input type="checkbox"/> Yes <input type="checkbox"/> No <input type="checkbox"/> Partially <input type="checkbox"/> Not applicable</p> <p><i>Coding guide:</i> YES: Both criteria should be fulfilled (where applicable) NO: Criteria not fulfilled PARTIALLY: Only one criteria fulfilled, or when there is limited reporting of quality appraisal (the latter applies only when inclusion criteria for study design are appropriate) NOT APPLICABLE: No included studies</p> <p><i>Note on reporting evidence and risk of bias:</i> For reviews of effects of 'large n' interventions, experimental and quasi-experimental designs should be included (if available). For reviews of effects of 'small n' interventions, designs appropriate to attribute changes to the intervention should be included (e.g. pre-post with assessment of confounders)</p>
<i>Please specify included study designs and any other comments (note important limitations or uncertainty):</i>	
<p>B.6 Did the review examine the extent to which specific factors might explain differences in the results of the included studies?</p> <p><input type="checkbox"/> Were factors that the review authors considered as likely explanatory factors clearly described?</p> <p><input type="checkbox"/> Was a sensible method used to explore the extent to which key factors explained heterogeneity?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Descriptive/textual <input type="checkbox"/> Graphical <input type="checkbox"/> Meta-analysis by sub-groups <input type="checkbox"/> Meta-regression <input type="checkbox"/> Other 	<p><input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable</p> <p><i>Coding guide:</i> YES: Explanatory factors clearly described and appropriate methods used to explore heterogeneity PARTIALLY: Explanatory factors described but for meta-analyses, sub-group analysis or meta-regression not reported (when they should have been) NO: No description or analysis of likely explanatory factors NOT APPLICABLE: e.g. too few studies, no important differences in the results of the included studies, or the included studies were so dissimilar that it would not make sense to explore heterogeneity of the results</p>
<i>Comments (note important limitations or uncertainty)</i>	

Section C: Overall assessment of the reliability of the review

<p>C.1 Are there any other aspects of the review not mentioned before which lead you to question the results?</p>	<p><input type="checkbox"/> Additional methodological concerns – only one person reviewing</p> <p><input type="checkbox"/> Robustness</p> <p><input type="checkbox"/> Interpretation</p> <p><input type="checkbox"/> Conflicts of interest (of the review authors or for included studies)</p> <p><input type="checkbox"/> Other</p> <p><input type="checkbox"/> No other quality issues identified</p>
<p>C.2 Are there any mitigating factors which should be taken into account in determining the reviews reliability?</p>	<p><input type="checkbox"/> Limitations acknowledged</p> <p><input type="checkbox"/> No strong policy conclusions drawn (including in abstract/ summary)</p> <p><input type="checkbox"/> Any other factors</p>
<p><i>Use comments to specify if relevant, to flag uncertainty or need for discussion</i></p>	
<p>C.3 Based on the above assessments of the methods please provide a summary of the quality of the review</p> <p>Strengths and limitations should be summarised above, based on what was noted in Sections A, B and C.</p>	

NOTES

ⁱ Adapted from Supporting the Use of Research Evidence (SURE) Collaboration. SURE checklist for making judgements about how much confidence to place in a systematic review. In: SURE guides for preparing and using policy briefs. www.evipnet.org/sure

ⁱⁱ **Risk of bias** is the extent to which bias may be responsible for the findings of a study.

Bias is a systematic error or deviation from the truth in results or inferences. In studies of the effects of social, economic and health care interventions, the main types of bias arise from systematic differences in the groups that are compared (selection bias), the intervention that is provided, or exposure to other factors apart from the intervention of interest (performance bias/contamination), withdrawals or exclusions of people entered into a study (attrition bias) or how outcomes are assessed (detection bias) and reported (reporting bias). Reviews of social science studies may be particularly affected by reporting bias, where a biased subset of all the relevant data and analyses is presented.

Assessments of the risk of bias are sometimes also referred to as assessments of the **validity** or **quality** of a study.

Validity is the extent to which a result (of a measurement or study) is likely to be true.

Quality is a vague notion of the strength or validity of a study, often indicating the extent of control over bias.