6 ทศวรรษรามาธิบดี
สู่สถาบันการแพทย์ในระดับสากล

คณะแพทยศาสตร์โรงพยาบาลรามาธิบดี
FACULTY OF MEDICINE RAMATHIBODI HOSPITAL

Cochrane
Evidence Synthesis and Methods

Ten Tips for Successful Assessment of Risk of Bias in Randomized Trials Using the RoB 2 Tool: Early lessons from Cochrane

Theresa H. M. Moore, Julian P. T. Higgins, Kerry Dwan

Commentator: Sureerat Suwatcharangkoon, MD

INTRODUCTION

- RoB 2 is a tool used by systematic reviewers to assess risk of bias in randomized trials.
- Over the period from June 2019 to December 2021 editors (KD and THMM) in the Cochrane Methods Support Unit peer reviewed 144 reviews, we saw many instances where users of RoB 2 frequently applied the tool in ways the developers had not intended, despite availability of detailed guidance, webinars and FAQs.

Methods

- In this paper we highlight the ten main issues that we observed, with the aims of optimising the application of the RoB 2 tool, avoiding some of the frequent misapplications of the tool and demonstrating how to present RoB 2 judgments within a review.

Structure of the RoB 2 tool for Assessing Bias in RCTs

Feature	About the tool	Comments
Focus of assessment	Results of randomized trials Results of "quasi-randomized" trials in which allocation was by means other than, but similar to, randomization (e.g., days of the week, birthdate, and so on).	Specific numerical results are assessed. If there is no numerical result for an outcome from a specific study, then there is no need to complete a RoB assessment as it will not contribute to a quantitative synthesis.
Effect of interest	Effect of assignment to intervention or Effect of adhering to a defined intervention	Reviews assessing the effects of an intervention will overwhelmingly be assessing the effect of assignment. The effect of adhering to an intervention can be useful for interventions for adverse events or to take the perspective of the health care user.
Five domains	1. Bias arising from the randomization process 2. Bias due to deviations from intended interventions 3. Bias due to missing outcome data 4. Bias in measurement of the outcome 5. Bias in selection of the reported result	All five domains should be assessed for all trials.

Structure of the RoB 2 tool for Assessing Bias in RCTs

Feature	About the tool	Comments
Level of bias	Each domain can be assessed as having: • Low risk of bias, • Some concerns about risk of bias, or • High risk of bias.	The judgment of risk of bias is determined from the answers to a series of "signaling questions" about the trial's conduct and course. An algorithm processes those answers into one of the three judgments.
Signaling questions	Between two and six signaling questions are used to inform the judgments about risk of bias. Answers to signaling questions are • Yes/Probably yes • No/Probably no • No information	The answers to signaling questions are recorded together with a brief reason for the answer.
Algorithm	An algorithm is built into the tool, to enable consistent choice of risk of bias for each domain.	
Overall risk of bias	Overall risk of bias is determined by considering the risks of bias in each of the five domains	• Low risk of bias: All domains are Low risk of bias • Some concerns: At least one domain is Some concerns and none are High risk of bias • High risk of bias: Any single domain is High risk of bias • [optional over-ride] High risk of bias: several domains are Some concerns such that in combination they warrant a judgment of High risk of bias

ASPECTS TO PLAN IN ADVANCE

1. Do plan assessments in advance
2. Do state the effect of interest
3. Do pilot the tool to reduce inconsistency in judgments

Preliminary considerations

Study design

- ☐ Individually-randomized parallel-group trial
- ☐ Cluster-randomized parallel-group trial
- ☐ Individually randomized cross-over (or other matched) trial

For the purposes of this assessment, the interventions being compared are defined as

Experimental: Comparator:

Specify which outcome is being assessed for risk of bias

Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Is the review team's aim for this result…?

- ☐ to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)
- ☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

If the aim is to assess the effect of *adhering to intervention*, select the deviations from intended intervention that should be addressed (at least one must be checked):

- ☐ occurrence of non-protocol interventions
- ☐ failures in implementing the intervention that could have affected the outcome
- ☐ non-adherence to their assigned intervention by trial participants

Tip 1: Do plan assessments in advance

- State the outcomes (including measures and timepoints) that will be addressed.
- RoB 2 may be applied separately to all outcomes, or to a subset of outcomes most important to decision makers.
- Be clear what time points and measurement methods are eligible for each synthesis within the review, because this will affect the answers to signaling questions in Domain 5 "bias in the selection of the reported result".

Preliminary considerations

Study design

- ☐ Individually-randomized parallel-group trial
- ☐ Cluster-randomized parallel-group trial
- ☐ Individually randomized cross-over (or other matched) trial

For the purposes of this assessment, the interventions being compared are defined as

Experimental: Comparator:

Specify which outcome is being assessed for risk of bias

Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Is the review team's aim for this result…?

- ☐ to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)
- ☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

If the aim is to assess the effect of *adhering to intervention*, select the deviations from intended intervention that should be addressed (at least one must be checked):

- ☐ occurrence of non-protocol interventions
- ☐ failures in implementing the intervention that could have affected the outcome
- ☐ non-adherence to their assigned intervention by trial participants

Tip 2: Do state the effect of interest

- Choose the effect of interest: either the effect of assignment to intervention or the effect of adhering to intervention.
- The signaling questions asked in Domain 2 "Bias due to deviations from the intended intervention" are affected by this decision.

Preliminary considerations

Study design

- ☐ Individually-randomized parallel-group trial
- ☐ Cluster-randomized parallel-group trial
- ☐ Individually randomized cross-over (or other matched) trial

For the purposes of this assessment, the interventions being compared are def

Experimental: Comparator:

Specify which outcome is being assessed for risk of bias

Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Is the review team's aim for this result…?

- ☐ to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)
- ☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

If the aim is to assess the effect of *adhering to intervention*, select the deviations from intended intervention that should be addressed (at least one must be checked):

- ☐ occurrence of non-protocol interventions
- ☐ failures in implementing the intervention that could have affected the outcome
- ☐ non-adherence to their assigned intervention by trial participants

Tip 3: Do pilot the tool to reduce inconsistency in judgments

- Develop a review-specific guidance document to help the team interpret the generic guidance for the specific topic under review.
- Pilot use of the RoB 2 tool for a few trials and discuss discrepancies to inform this document, to help ensure consistent assessments.
- Inexperienced users of RoB 2 are paired with more experienced users to do independent assessments.

FACTORS TO CONSIDER WHEN APPLYING THE TOOL

4. Do apply the tool to a specific numerical result and not the whole study
5. Do answer all signaling questions, use the algorithm and provide supporting information for judgments

Tip 4: Do apply the tool to a specific numerical result and not the whole study

- Risk of bias may differ for different outcomes, and even for different results for the same outcome.
- Avoid attempting to apply the tool to a trial as a whole.

Preliminary considerations

Study design

- ☐ Individually-randomized parallel-group trial
- ☐ Cluster-randomized parallel-group trial
- ☐ Individually randomized cross-over (or other matched) trial

For the purposes of this assessment, the interventions being compared are def

Experimental: | Comparator:

Specify which outcome is being assessed for risk of bias

Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Is the review team's aim for this result…?

- ☐ to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)
- ☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

If the aim is to assess the effect of *adhering to intervention*, select the deviations from intended intervention that should be addressed (at least one must be checked):

- ☐ occurrence of non-protocol interventions
- ☐ failures in implementing the intervention that could have affected the outcome
- ☐ non-adherence to their assigned intervention by trial participants

Preliminary considerations

Study design

- ☐ Individually-randomized parallel-group trial
- ☐ Cluster-randomized parallel-group trial
- ☐ Individually randomized cross-over (or other matched) trial

For the purposes of this assessment, the interventions being compared are def

Experimental: | Comparator:

Specify which outcome is being assessed for risk of bias

Specify the numerical result being assessed. In case of multiple alternative analyses being presented, specify the numeric result (e.g. RR = 1.52 (95% CI 0.83 to 2.77) and/or a reference (e.g. to a table, figure or paragraph) that uniquely defines the result being assessed.

Is the review team's aim for this result…?

- ☐ to assess the effect of *assignment to intervention* (the 'intention-to-treat' effect)
- ☐ to assess the effect of *adhering to intervention* (the 'per-protocol' effect)

If the aim is to assess the effect of *adhering to intervention*, select the deviations from intended intervention that should be addressed (at least one must be checked):

- ☐ occurrence of non-protocol interventions
- ☐ failures in implementing the intervention that could have affected the outcome
- ☐ non-adherence to their assigned intervention by trial participants

Tip 5: Do answer all signaling questions, use the algorithm and provide supporting information for judgments

- The RoB 2 tool asks users to complete all signaling questions, provide support for judgments, and use the algorithm to make the bias assessment.
- Omission of any of these steps can lead to inaccurate assessments (overly harsh or overly lenient) and a lack of transparency.

COMMON PROBLEMS WITH SPECIFIC DOMAINS

6. Don't assume baseline imbalance necessarily means bias (Domain1)
7. Don't assume no blinding means bias (Domains 2 and 4)
8. Don't assume switching interventions necessarily means bias (Domain 2)
9. Don't set arbitrary thresholds for missing outcome data (Domain 3)
10. Don't assume the absence of a statistical analysis plan means bias (Domain 5)

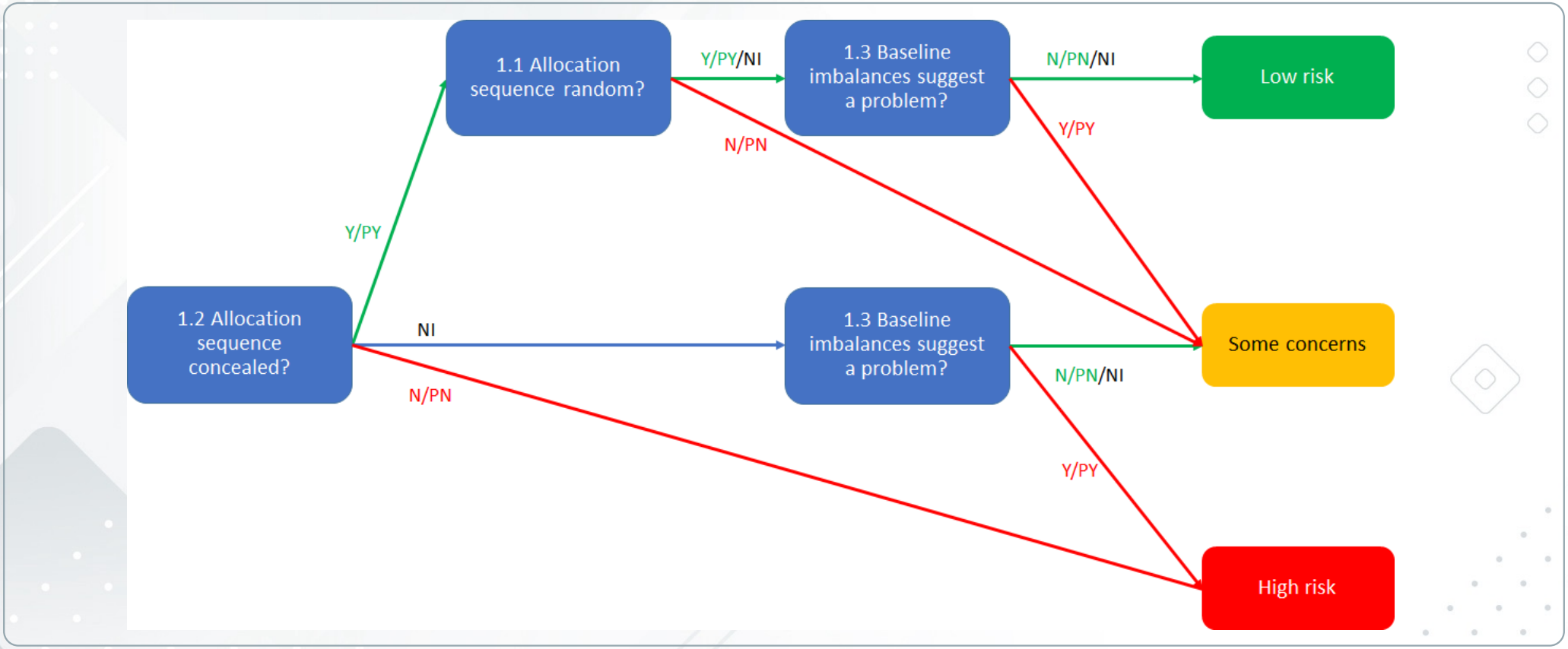
Domain 1: Risk of Bias Arising from the Randomization Process

Signalling questions	Elaboration	Response options
1.3 Did baseline differences between intervention groups suggest a problem with the randomization process?	*Note that differences that are compatible with chance do not lead to a risk of bias. A small number of differences identified as 'statistically significant' at the conventional 0.05 threshold should usually be considered to be compatible with chance.* Answer 'No' if no imbalances are apparent or if any observed imbalances are compatible with chance. Answer 'Yes' if there are imbalances that indicate problems with the randomization process, including: (1) substantial differences between intervention group sizes, compared with the intended allocation ratio; or (2) a substantial excess in statistically significant differences in baseline characteristics between intervention groups, beyond that expected by chance; or (3) imbalance in one or more key prognostic factors, or baseline measures of outcome variables, that is very unlikely to be due to chance and for which the between-group difference is big enough to result in bias in the intervention effect estimate. Also answer 'Yes' if there are other reasons to suspect that the randomization process was problematic: (4) excessive similarity in baseline characteristics that is not compatible with chance. Answer 'No information' when there is no *useful* baseline information available (e.g. abstracts, or studies that reported only baseline characteristics of participants in the final analysis). The answer to this question should not influence answers to questions 1.1 or 1.2. For example, if the trial has large baseline imbalances, but authors report adequate randomization methods, questions 1.1 and 1.2 should still be answered on the basis of the reported adequate methods, and any concerns about the imbalance should be raised in the answer to the question 1.3 and reflected in the domain-level risk-of-bias judgement. Trialists may undertake analyses that attempt to deal with flawed randomization by controlling for imbalances in prognostic factors at baseline. To remove the risk of bias caused by problems in the randomization process, it would be necessary to know, and measure, all the prognostic factors that were imbalanced at baseline. It is unlikely that all important prognostic factors are known and measured, so such analyses will at best reduce the risk of bias. If review authors wish to assess the risk of bias in a trial that controlled for baseline imbalances in order to mitigate failures of randomization, the study should be assessed using the ROBINS-I tool.	Y/PY/PN/N/NI
Risk-of-bias judgement	See algorithm.	Low / High / Some concerns

Tip 6: Don't assume baseline imbalance necessarily means bias (Domain 1)

- Some degree of baseline imbalance is expected to occur by chance in any randomized trial.
- It is important to consider whether the imbalance is sufficiently extreme to indicate that something has gone wrong with the randomization process.

Algorithm for Suggested Judgement of Risk of Bias arising from the Randomization Process



Domain 2: Risk of Bias due to Deviations from the Intended Interventions (Effect of Assignment to Intervention)

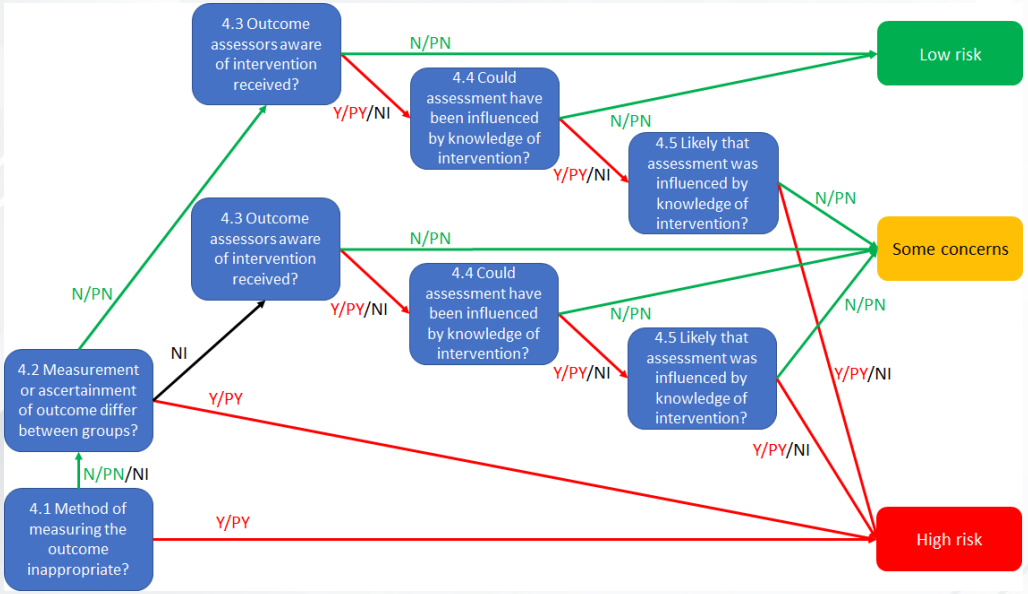
Signalling questions	Elaboration	Response options
2.1. Were participants aware of their assigned intervention during the trial?	If participants are aware of their assigned intervention it is more likely that health-related behaviours will differ between the intervention groups. Blinding participants, most commonly through use of a placebo or sham intervention, may prevent such differences. If participants experienced side effects or toxicities that they knew to be specific to one of the interventions, answer this question 'Yes' or 'Probably yes'.	Y/PY/PN/N/NI
2.2. Were carers and people delivering the interventions aware of participants' assigned intervention during the trial?	If carers or people delivering the interventions are aware of the assigned intervention then its implementation, or administration of non-protocol interventions, may differ between the intervention groups. Blinding may prevent such differences. If participants experienced side effects or toxicities that carers or people delivering the interventions knew to be specific to one of the interventions, answer question 'Yes' or 'Probably yes'. If randomized allocation was not concealed, then it is likely that carers and people delivering the interventions were aware of participants' assigned intervention during the trial.	
2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?	For the effect of assignment to intervention, this domain assesses problems that arise when changes from assigned intervention that are inconsistent with the trial protocol arose because of the trial context. We use the term **trial context** to refer to effects of recruitment and engagement activities on trial participants and when trial personnel (carers or people delivering the interventions) undermine the implementation of the trial protocol in ways that would not happen outside the trial. For example, the process of securing informed consent may lead participants subsequently assigned to the comparator group to feel unlucky and therefore seek the experimental intervention, or other interventions that improve their prognosis. Answer 'Yes' or 'Probably yes' **only** if there is evidence, or strong reason to believe, that the trial context led to failure to implement the protocol interventions or to implementation of interventions not allowed by the protocol. Answer 'No' or 'Probably no' if there were changes from assigned intervention that are inconsistent with the trial protocol, such as non-adherence to intervention, but these are consistent with what could occur outside the trial context. Answer 'No' or 'Probably no' for changes to intervention that are consistent with the trial protocol, for example cessation of a drug intervention because of acute toxicity or use of additional interventions whose aim is to treat consequences of one of the intended interventions. If blinding is compromised because participants report side effects or toxicities that are specific to one of the interventions, answer 'Yes' or 'Probably yes' only if there were changes from assigned intervention that are inconsistent with the trial protocol and arose because of the trial context. The answer 'No information' may be appropriate, because trialists do not always report whether deviations arose because of the trial context.	

Tip 7: Don't assume no blinding means bias (Domains 2 and 4)

- Lack of blinding of participants or outcome assessors does not always indicate bias.
- Randomized trials can be open label yet not troubled by bias.
- Users should take time to read the guidance and answer all the signaling questions to ensure they consider whether knowing the assignment was likely to lead to bias.

Domain 4: Risk of bias in measurement of the outcome

Signalling questions	Elaboration	Response options
4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?	Answer 'No' if outcome assessors were blinded to intervention status. For participant-reported outcomes, the outcome assessor is the study participant.	NA/Y/PY/PN/N/NI



Tip 7: Don't assume no blinding means bias (Domains 2 and 4)

- We observed a similar issue relating to outcome assessment: it was common for a "High risk of bias" judgment to be reached purely on the basis that outcome assessors were aware of intervention received.
- Subsequent signaling questions exploring whether this lack of blinding would impact on assessments of the outcome (and hence lead to a risk of bias) were ignored, and so the algorithm was not used.

Domain 2: Risk of bias due to deviations from the intended interventions (effect of adhering to intervention)

Signalling questions	Elaboration	Response options
2.5. [If applicable:] Was there non-adherence to the assigned intervention regimen that could have affected participants' outcomes?	This question is asked only if the preliminary considerations specify that the assessment will address non-adherence that could have affected participants' outcomes. Non-adherence includes imperfect compliance with a sustained intervention, cessation of intervention, crossovers to the comparator intervention and switches to another active intervention. Consider available information on the proportion of study participants who continued with their assigned intervention throughout follow up, and answer 'Yes' or 'Probably yes' if the proportion who did not adhere is high enough to raise concerns. Answer 'No' for studies of interventions that are administered once, so that imperfect adherence is not possible, and all or most participants received the assigned intervention.	NA/Y/PY/PN/N/NI

Tip 8: Don't assume switching interventions necessarily means bias (Domain 2)

- Subsequent signaling questions exploring whether this lack of blinding would impact on assessments of the outcome (and hence lead to a risk of bias) were ignored, and so the algorithm was not used.
- Not all changes in treatment delivered within a trial present a risk of bias.
- For example, clinicians may change a treatment strategy if a participant's disease progresses.
- Changes that would occur outside of the trial context, such as due to disease progression, do not introduce bias in the effect of assignment to intervention.
- Problems arise only if changes in intervention by trial participants happen because of the trial context.
- We advise authors to check the definition of a "Deviation of intervention" in the detailed guidance and discuss with the review team what would be classed as a deviation from intervention.
- Draft a RoB 2 consensus document for the review team that includes the definition and examples.

Domain 3: Risk of bias due to missing outcome data

Signalling questions	Elaboration	Response options
3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?	Evidence that the result was not biased by missing outcome data may come from: (1) analysis methods that correct for bias; or (2) sensitivity analyses showing that results are little changed under a range of plausible assumptions about the relationship between missingness in the outcome and its true value. However, imputing the outcome variable, either through methods such as 'last-observation-carried-forward' or via multiple imputation based only on intervention group, should not be assumed to correct for bias due to missing outcome data.	NA/Y/PY/PN/N
3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?	If loss to follow up, or withdrawal from the study, could be related to participants' health status, then it is possible that missingness in the outcome was influenced by its true value. However, if all missing outcome data occurred for documented reasons that are unrelated to the outcome then the risk of bias due to missing outcome data will be low (for example, failure of a measuring device or interruptions to routine data collection). In time-to-event analyses, participants censored during trial follow-up, for example because they withdrew from the study, should be regarded as having missing outcome data, even though some of their follow up is included in the analysis. Note that such participants may be shown as included in analyses in CONSORT flow diagrams.	
3.4 If Y/PY/NI to 3.3: Is it likely that missingness in the outcome depended on its true value?	This question distinguishes between situations in which (i) missingness in the outcome could depend on its true value (assessed as 'Some concerns') from those in which (ii) it is likely that missingness in the outcome depended on its true value (assessed as 'High risk of bias'). Five reasons for answering 'Yes' are: 1. Differences between intervention groups in the proportions of missing outcome data. If there is a difference between the effects of the experimental and comparator interventions on the outcome, and the missingness in the outcome is influenced by its true value, then the proportions of missing outcome data are likely to differ between intervention groups. Such a difference suggests a risk of bias due to missing outcome data, because the trial result will be sensitive to missingness in the outcome being related to its true value. For time-to-event-data, the analogue is that rates of censoring (loss to follow-up) differ between the intervention groups. 2. Reported reasons for missing outcome data provide evidence that missingness in the outcome depends on its true value; 3. Reported reasons for missing outcome data differ between the intervention groups; 4. The circumstances of the trial make it likely that missingness in the outcome depends on its true value. For example, in trials of interventions to treat schizophrenia it is widely understood that continuing symptoms make drop out more likely. 5. In time-to-event analyses, participants' follow up is censored when they stop or change their assigned intervention, for example because of drug toxicity or, in cancer trials, when participants switch to second-line chemotherapy. Answer 'No' if the analysis accounted for participant characteristics that are likely to explain the relationship between missingness in the outcome and its true value.	

Tip 9: Don't set arbitrary thresholds for missing outcome data (Domain 3)

- Avoid setting an arbitrary threshold for assessing the amount of missing outcome data.
- Ensure that all relevant signaling questions are answered. See the detailed guidance for more information.

Domain 5: Risk of bias in selection of the reported result

Domain 5: Risk of bias in selection of the reported result

Signalling questions	Elaboration
5.1 Were the data that produced this result analysed in accordance with a pre-specified analysis plan that was finalized before unblinded outcome data were available for analysis?	If the researchers' pre-specified intentions are available in sufficient detail, then planned outcome measurements and analyses can be compared with those presented in the published report(s). To avoid the possibility of selection of the reported result, finalization of the analysis intentions must precede availability of unblinded outcome data to the trial investigators. Changes to analysis plans that were made before unblinded outcome data were available, or that were clearly unrelated to the results (e.g. due to a broken machine making data collection impossible) do not raise concerns about bias in selection of the reported result.
Is the numerical result being assessed likely to have been selected, on the basis of the results, from...	
5.2. ... multiple eligible outcome measurements (e.g. scales, definitions, time points) within the outcome domain?	A particular outcome domain (i.e. a true state or endpoint of interest) may be **measured** in multiple ways. For example, the domain pain may be measured using multiple scales (e.g. a visual analogue scale and the McGill Pain Questionnaire), each at multiple time points (e.g. 3, 6 and 12 weeks post-treatment). If multiple measurements were made, but only one or a subset is reported on the basis of the results (e.g. statistical significance), there is a high risk of bias in the fully reported result. Attention should be restricted to outcome measurements that are eligible for consideration by the RoB 2 tool user. For example, if only a result using a specific measurement scale is eligible for inclusion in a meta-analysis (e.g. Hamilton Depression Rating Scale), and this is reported by the trial, then there would not be an issue of selection even if this result was reported (on the basis of the results) in preference to the result from a different measurement scale (e.g. Beck Depression Inventory). Answer 'Yes' or 'Probably yes' if: There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that a domain was measured in multiple eligible ways, but data for only one or a subset of measures is fully reported (without justification), and the fully reported result is likely to have been selected on the basis of the results. Selection on the basis of the results can arise from a desire for findings to be newsworthy, sufficiently noteworthy to merit publication, or to confirm a prior hypothesis. For example, trialists who have a preconception, or vested interest in showing, that an

Tip 10: Don't assume the absence of a statistical analysis plan means bias (Domain 5)

- A protocol or a trial registration document or a statistical analysis plan can be used to address risk of bias for this domain.
- It is <u>not</u> necessary to find the formal statistical analysis plan for a randomized trials as frequently these are <u>not</u> available.
- Some protocols or plans are, unfortunately, registered retrospectively.
- Therefore it is important to check that registration was before data analysis.

DISCUSSION

- We observed a combination of misapplications of the tool, which can lead to overly harsh or overly lenient assessments of bias, and the lack of transparency can lead to lack of confidence from review users in the contents of the review.
- Poor reporting was particularly notable
  - for the effect of interest
  - for which outcomes were to be assessed
  - for the rationale for judgments of risk of bias

Reason 1
Complex Process

- Difficult to apply and time consuming to use, even by experienced reviewers.
- It requires users to be familiar with the methods of conducting randomized trials, statistical analysis of randomized trials, and the nature and implications of variation in how interventions are implemented in practice.

Reason 1
Complex Process

- In the initial Cochrane risk-of-bias tool, no distinction was made between interest in the effect of assignment to intervention and the effect of adhering to intervention, meaning that assessments of risk of bias due to lack of blinding in an open-label trial were unfocussed.
- In RoB 2, the distinction allows for open-label trials to be judged, appropriately, to be at low risk of bias due to deviations from intended interventions.

Reason 2
Lack of Time to Learn the New Tool

- The detailed RoB 2 guidance document is lengthy, and some key aspects might not be immediately obvious to users.
- Piloting the bias tool could also improve consistency in judgments, as has been seen in the use of the ROBINS-I tool for assessing risk of bias in non-randomized studies.
- We recommend that systematic review teams using the RoB 2 tool have at least one member who is fully familiar with the tool and detailed guidance, possibly making use of a series of webinars.

Guidance and training resources for users of risk of bias 2

CONCLUSIONS

- RoB 2 has been used inappropriately in systematic reviews submitted for editorial peer review.
- Clear and transparent reporting of research is essential for reproducibility and replication.
- Resources are available to assist users of the tool and strategies that might help include becoming familiar with the detailed guidance, piloting the tool as a team to reduce inconsistency, engaging with available learning resources
  - Introduction to RoB 2; Cochrane Starter pack; FAQs; webinar series; detailed RoB 2 guidance; and a Checklist for editors and peer reviewers