RoB 2: a revised tool for assessing risk of bias in randomised trials

Jonathan A C Sterne,1,2 Jelena Savović,1,3 Matthew J Page,4 Roy G Elbers,3 Natalie S Blencowe,1,2 Isabelle Bouton,5,6,7 Christopher J Cates,8 Hung-Yuan Cheng,1,2 Mark S Corbett,9 Sandra M Eldridge,10 Jonathan R Emberson,11 Miguel A Hernán,12 Sally Hopewell,13 Asbjørn Hróbjartsson,14,15,16 Daniela R Junqueira,17 Peter Juni,18 Jamie J Kirkham,19 Toby Lasserson,20 Tianjing Li,21 Alexandra McAleenan,1 Barnaby C Reeves,2,22 Sandra M Eldridge,10 Jonathan R Emberson,11 Miguel A Hernán,12 Sally Hopewell,13 Asbjørn Hróbjartsson,14,15,16 Daniela R Junqueira,17 Peter Juni,18 Jamie J Kirkham,19 Toby Lasserson,20 Tianjing Li,21

Assessment of risk of bias is regarded as an essential component of a systematic review on the effects of an intervention. The most commonly used tool for randomised trials is the Cochrane risk-of-bias tool. We updated the tool to respond to developments in understanding how bias arises in randomised trials, and to address user feedback on and limitations of the original tool.

An evaluation of the risk of bias in each study included in a systematic review documents potential flaws in the evidence summarised and contributes to the certainty in the overall evidence.1 The Cochrane tool for assessing risk of bias in randomised trials (RoB tool)2 has been widely used in both Cochrane and other systematic reviews, with over 40 000 citations in Google Scholar.

Many innovative characteristics of the original RoB tool have been widely accepted. It replaced the notion of assessing study quality with that of assessing risk of bias (we define bias as a systematic deviation from the effect of intervention that would be observed in a large randomised trial without any flaws). Quality is not well defined and can include study characteristics (such as performing a sample size calculation) that are not inherently related to bias in the study’s results. The RoB tool considers biases arising at different stages of a trial (known as bias domains), which were chosen on the basis of both empirical evidence and theoretical considerations. Assessments of risk of bias are supported by quotes from sources describing the trial (eg, trial protocol, registration record, results report) or by justifications written by the assessor.

After nearly a decade of experience of using the RoB tool, potential improvements have been identified. A formal evaluation found some bias domains to be confusing at times, with assessment of bias due to incomplete outcome data and selective reporting of outcomes causing particular difficulties, and confusion over whether studies that were not blinded should automatically be considered to be at high risk of bias.3 More guidance on incorporating risk-of-bias assessments into meta-analyses and review conclusions is also needed.4 5 A review of comments and user practice found that both Cochrane and non-Cochrane systematic reviews often implemented the RoB tool in non-standard ways.6 Few trials are assessed as at low risk of bias, and judgments of unclear risk of bias are common.6,7 Empirical studies have found only moderate reliability of risk-of-bias judgments.8

We developed a revised risk-of-bias assessment tool to address these issues, incorporate advances in assessment of risk of bias used in other recently developed tools,9 10 and integrate recent developments in estimation of intervention effects from randomised trials.11

Development of the revised RoB tool

We followed the principles adopted for the development of the original RoB tool and for the ROBINS-I tool for assessing risk of bias in non-randomised studies of interventions.2 9 A core group coordinated development of the tool, including recruitment of collaborators, preparation and revision of documents, and administrative support.

Preliminary work included a review of how the original tool was used in practice,4 a systematic review and meta-analysis of meta-epidemiological studies of empirical evidence for biases associated

For numbered affiliations see end of the article.

Correspondence to: J A C Sterne jonathan.sterne@bristol.ac.uk (or @jonathanasterne on Twitter; ORCID 0000-0001-8496-6053)

Additional material is published online only. To view please visit the journal online.

Cite this as: BMJ 2019;366:l4898
http://dx.doi.org/10.1136/bmj.l4898

Accepted: 25 June 2019

SUMMARY POINTS

- Assessment of risk of bias is regarded as an essential component of a systematic review on the effects of an intervention; the most commonly used tool for assessing risk of bias in randomised trials is the Cochrane risk-of-bias tool, which was introduced in 2008.
- Potential improvements to the Cochrane risk-of-bias tool were identified on the basis of reviews of the literature, user experience and feedback, approaches used in other risk-of-bias tools, and recent developments in estimation of intervention effects from randomised trials.
- We developed and piloted a revised tool for assessing risk of bias in randomised trials (RoB 2).
- Bias is assessed in five distinct domains. Within each domain, users of RoB 2 answer one or more signalling questions. These answers lead to judgments of “low risk of bias,” “some concerns,” or “high risk of bias.”
- The judgments within each domain lead to an overall risk-of-bias judgment for the result being assessed, which should enable users of RoB 2 to stratify meta-analyses according to risk of bias.
with characteristics of randomised trials,\(^1\) and a cross-sectional study of how selective outcome reporting was assessed in Cochrane reviews.\(^1\) We also drew on a systematic review of the theoretical and conceptual literature on types of bias in epidemiology, which sought papers and textbooks presenting classifications or definitions of biases, and organised these into a coherent framework (paper in preparation).

The core group developed an initial proposal and presented it, together with the latest empirical evidence of biases in randomised trials, at a meeting in August 2015 attended by 24 contributors. Meeting participants agreed on the methodological principles underpinning the new tool and the bias domains to be addressed, and formed working groups for each domain. The groups were tasked with developing signalling questions (reasonably factual questions with yes/no answers that inform risk-of-bias judgments), together with guidance for answering these questions and broad considerations for how to judge the risk of bias for the domain.

The materials prepared by the working groups were assembled and edited by the core team, and the resulting draft was piloted by experienced and novice systematic reviewers during a three day event in February 2016, with 17 participants present and 10 participants contributing remotely. Issues identified in the pilot were recorded and addressed in a new draft discussed at a second development meeting in April 2016, also attended by 24 contributors. Subsequently, working groups developed criteria for reaching domain level, risk-of-bias judgments based on answers to signalling questions, and expanded the guidance.

The core team designed algorithms to match the criteria, which were checked by the working groups. The resulting revision was tested in another round of piloting by 10 systematic review authors in mid-2016.

A complete draft of version 2 of the RoB tool (RoB 2), together with detailed guidance, was posted at www.riskofbias.info in October 2016, coinciding with the Cochrane Colloquium in Seoul, South Korea. Feedback was invited through direct contact with the development group. Several review teams subsequently piloted the draft tool and provided feedback. Further modifications—particularly improvements in wording and clarity, splitting compound signalling questions, adding new questions, and addressing methodological issues—were made on the basis of feedback from training events (including webinars) conducted between 2016 and 2019, as well as individual feedback from users worldwide.

### Version 2 of the Cochrane tool for assessing risk of bias in randomised trials (RoB 2)

RoB 2 provides a framework for assessing the risk of bias in a single estimate of an intervention effect reported from a randomised trial. The effect assessed is a comparison of two interventions, which we refer to as the experimental and comparator interventions, for a specific outcome or endpoint. The process of making a RoB 2 assessment is summarised in figure 1. Preliminary considerations (box 1) include specifying which result is being assessed, specifying how this result is being interpreted (see “The intervention effect of interest” below), and listing the sources of information used to inform the assessment. Review authors should contact trial authors in order to obtain information that is omitted from published and online sources, so far as this is feasible. Note that risk-of-bias assessments might be needed for results relating to multiple outcomes from the included trials.

RoB 2 is structured into five bias domains, listed in table 1. The domains were selected to address all important mechanisms by which bias can be introduced into the results of a trial, based on a combination of empirical evidence and theoretical considerations. We did not include domains for features that would be expected to operate indirectly, through the included bias domains.\(^14\)\(^\)\(^1\)\(^5\) For this reason, we excluded some trial features, such as funding source and single centre versus multicentre status, which have been associated empirically with trial effect estimates from trials.

We label the domains using descriptions of the causes of bias addressed, avoiding terms used in the original RoB tool (such as “selection bias” and “performance bias”) because they are used inconsistently or not known by many people outside Cochrane.\(^1\)\(^6\) Each domain is mandatory, and no others can be added, although we have developed versions of RoB 2 that deal with additional issues that arise in trials with cluster randomised or crossover designs (www.riskofbias.info). Within each domain, the assessment comprises:

<table>
<thead>
<tr>
<th>For each outcome</th>
<th>For each study</th>
<th>For the synthesis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk of bias assessment for a specific result</td>
<td>1. Specify result being assessed</td>
<td>Integrate judgment(s) into results and conclusions</td>
</tr>
<tr>
<td>2. Specify effect of interest</td>
<td>3. List sources of information used to inform assessment</td>
<td>Eg. stratify meta-analysis by overall risk-of-bias judgment</td>
</tr>
<tr>
<td>3. Answer signalling questions</td>
<td>4. Judge risk of bias for each domain</td>
<td></td>
</tr>
<tr>
<td>5. Judge risk of bias for each domain</td>
<td>6. Judge overall risk of bias for the result</td>
<td></td>
</tr>
<tr>
<td>6. Judge overall risk of bias for the result</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Fig 1 | Summary of the process of assessing risk of bias in a systematic review of randomised trials, using version 2 of the Cochrane risk-of-bias tool
• A series of signalling questions
• A judgment about risk of bias for the domain, facilitated by an algorithm that maps responses to signalling questions to a proposed judgment
• Free text boxes to justify responses to the signalling questions and risk-of-bias judgments
• Optional free text boxes to predict (and explain) the likely direction of bias.

Table 2 lists the most important changes made in RoB 2, compared with the original Cochrane RoB tool.

Signalling questions
Signalling questions aim to elicit information relevant to an assessment of risk of bias (table 1). The questions seek to be reasonably factual in nature. The response options are “yes,” “probably yes,” “probably no,” “no,” and “no information.” To maximise their simplicity and clarity, signalling questions are phrased such that a yes answer might indicate either lower or higher risk of bias, depending on the most natural way to ask the question. The online supplementary material in the web appendix includes elaborations providing guidance on how to answer each question.

Responses of “yes” and “probably yes” have the same implications for risk of bias, as do responses of “no” and “probably no.” “Yes” and “no” typically imply that firm evidence is available; the “probably” responses typically imply that a judgment has been made. Where there is a need to distinguish between “some concerns” and “high risk of bias,” this is dealt with by using an additional signalling question, rather than by making a distinction between responses “probably yes” and “yes,” or between “probably no” and “no.” The “no information” response should be used only when insufficient details are available to allow a different response, and when, in the absence of these details, it would be unreasonable to respond “probably yes” or “probably no.” For example, in the context of a large trial run by an experienced clinical trials unit, absence of specific information about generation of the randomisation sequence, in a paper published in a journal with rigorously enforced word count limits, is likely to result in a response of “probably yes” rather than “no information” to the signalling question about sequence generation (the rationale for the response should be provided in the free text box). Some signalling questions are answered only if the response to a previous question indicates that they are required.

The intervention effect of interest
Assessments for the domain “bias due to deviations from intended interventions” differ according to whether review authors are interested in quantifying the effect of assignment to the interventions at baseline regardless of whether the interventions are received during follow-up (intention-to-treat effect), or the effect of adhering to intervention as specified in the trial protocol (per protocol effect). These effects will differ if some patients do not receive their assigned intervention or deviate from the assigned intervention after baseline. Each effect might be of interest.11 For example, the effect of assignment to intervention might be appropriate to inform a health policy question about whether to recommend an intervention (eg, a screening programme) in a particular health system, whereas the effect of adhering to intervention more directly informs a care decision by an individual patient (eg, whether to be screened). Changes to an intervention that are consistent with the trial protocol (even if not explicitly discussed in the protocol), such as cessation of a drug because of toxicity or switch to second line chemotherapy because of progression of cancer, do not cause bias and should not be considered to be deviations from intended intervention.

The effect of assignment to intervention should be estimated by an intention-to-treat analysis that includes all randomised participants.17 However, estimates of per protocol effects commonly used in reports of randomised trials are problematic and might be seriously biased.18 These estimates include those from naive per protocol analyses restricted to individuals who adhered to their assigned intervention, and as-treated analyses in which participants are analysed according to the intervention they received, even if their assigned group is different. These approaches are problematic because prognostic factors could

Box 1: RoB 2 tool: preliminary considerations

- For the purposes of this assessment, define the interventions being compared:
  - Experimental intervention:
  - Comparator intervention:
- 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 numerical result (eg, risk ratio 1.52 (95% confidence interval 0.83 to 2.77) or a reference (eg, to a table, figure, or paragraph) that uniquely defines the result being assessed.)
- Is the review team’s aim for this result (check one):
  - 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
- Which of the following sources were obtained to help inform the risk-of-bias assessment?
  - Journal article(s)
  - Trial protocol
  - Statistical analysis plan
  - Non-commercial trial registry record (eg, ClinicalTrials.gov record)
  - Company owned trial registry record (eg, GlaxoSmithKline Clinical Study Register record)
  - Grey literature (eg, unpublished thesis)
  - Conference abstract(s) about the trial
  - Regulatory document (eg, clinical study report, drug approval package)
  - Research ethics application
  - Grant database summary (eg, NIH RePORTER or Research Councils UK Gateway to Research)
  - Personal communication with triallist
  - Personal communication with the sponsor
influence whether individuals receive their allocated intervention. Data from a randomised trial can be used to derive an unbiased estimate of the effect of adhering to intervention. However, the validity of appropriate methods depends on strong assumptions, and published applications are relatively rare to date. For trials comparing interventions that are sustained over time, appropriate methods also require measurement of and adjustment for the values of prognostic factors, both before and after randomisation, that predict

<table>
<thead>
<tr>
<th>Bias domain and signalling question*</th>
<th>Response options</th>
<th>Lower risk of bias</th>
<th>Higher risk of bias</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias arising from the randomisation process</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1 Was the allocation sequence random?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>1.2 Was the allocation sequence concealed until participants were enrolled and assigned to interventions?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>1.3 Did baseline differences between intervention groups suggest a problem with the randomisation process?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td><strong>Risk-of-bias judgment (low/high/some concerns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias arising from the randomisation process?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bias due to deviations from intended interventions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Were participants aware of their assigned intervention during the trial?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>2.2 Were carers and people delivering the interventions aware of participants’ assigned intervention during the trial?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>2.3 If Y/PY/N to 2.1 or 2.2: Were there deviations from the intended intervention that arose because of the trial context?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td>2.4 If Y/PY/N to 2.3: Were these deviations likely to have affected the outcome?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td>2.5 If Y/PY to 2.4: Were these deviations from intended intervention balanced between groups?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td>2.6 Was an appropriate analysis used to estimate the effect of assignment to intervention?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>2.7 If N/PN/NI to 2.6: Was there potential for a substantial impact (on the result) of the failure to analyse participants in the group to which they were randomised?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td><strong>Risk-of-bias judgment (low/high/some concerns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias due to deviations from intended interventions?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bias due to missing outcome data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1 Were data for this outcome available for all, or nearly all, participants randomised?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>3.2 If N/PN/NI to 3.1: Is there evidence that the result was not biased by missing outcome data?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>3.3 If N/PN to 3.2: Could missingness in the outcome depend on its true value?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td>3.4 If Y/PY/N to 3.3: Is it likely that missingness in the outcome depended on its true value?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td><strong>Risk-of-bias judgment (low/high/some concerns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias due to missing outcome data?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bias in measurement of the outcome</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1 Was the method of measuring the outcome inappropriate?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>4.2 Could measurement or ascertainment of the outcome have differed between intervention groups?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>4.3 If N/PN/NI to 4.1 and 4.2: Were outcome assessors aware of the intervention received by study participants?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>4.4 If Y/PY/N to 4.3: Could assessment of the outcome have been influenced by knowledge of intervention received?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td>4.5 If Y/PY/N to 4.4: Is it likely that assessment of the outcome was influenced by knowledge of intervention received?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NA/NI</td>
<td></td>
</tr>
<tr>
<td><strong>Risk-of-bias judgment (low/high/some concerns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias in measurement of the outcome?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Bias in selection of the reported result</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.1 Were the data that produced this result analysed in accordance with a prespecified analysis plan that was finalised before unblinded outcome data were available for analysis?</td>
<td>Y/PY</td>
<td>N/PN</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>5.2 ... multiple eligible outcome measurements (eg, scales, definitions, time points) within the outcome domain?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td>5.3 ... multiple eligible analyses of the data?</td>
<td>N/PN</td>
<td>Y/PY</td>
<td>NI</td>
<td></td>
</tr>
<tr>
<td><strong>Risk-of-bias judgment (low/high/some concerns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction bias due to selection of the reported results?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall bias</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Risk-of-bias judgment (low/high/some concerns)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optional: What is the overall predicted direction of bias for this outcome?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Signalling questions for bias due to deviations from intended interventions relate to the effect of assignment to intervention.
Reasons for any judgments that do not follow the information obtained from authors when contacted. be used whenever possible, supplemented by any of the study reports (including trial protocols) should and judgments allow assessors to provide support for proposed judgments if they feel it is appropriate to each domain (see online supplementary material in questions to a proposed risk-of-bias judgment for of algorithms that map responses to signalling

---

**Risk-of-bias judgments**

The risk-of-bias judgments for each domain are “low risk of bias,” “some concerns,” or “high risk of bias.” Judgments are based on, and summarise, the answers to signalling questions. Review authors should interpret “risk of bias” as “risk of material bias”: concerns should be expressed only about issues likely to have a notable effect on the result being assessed.

An important innovation in RoB 2 is the inclusion of algorithms that map responses to signalling questions to a proposed risk-of-bias judgment for each domain (see online supplementary material in the web appendix). Review authors can override these proposed judgments if they feel it is appropriate to do so.

Free text boxes alongside the signalling questions and judgments allow assessors to provide support for the responses. Brief direct quotations from the texts of the study reports (including trial protocols) should be used whenever possible, supplemented by any information obtained from authors when contacted. Reasons for any judgments that do not follow the algorithms should be provided. RoB 2 includes optional judgments of the direction of the bias for each domain and overall. If review authors do not have a clear rationale for judging the likely direction of the bias, they should not guess it.

---

### Overall risk of bias for the result

The response options for an overall risk-of-bias judgment are the same as for individual domains. Table 3 shows the approach to mapping bias judgments within domains to an overall judgment for the result. The overall risk of bias generally corresponds to the worst risk of bias in any of the domains. However, if a study is judged to have “some concerns” about risk of bias for multiple domains, it might be judged as at high risk of bias overall. Figure 2 shows a forest plot that displays domain specific risk of bias and overall risk of bias, with the meta-analysis stratified by overall risk of bias.

### Discussion

We have substantially revised the Cochrane tool for assessing risk of bias in the results of randomised trials, in order to address limitations identified since it was published in 2008 and to incorporate improvements that aim to increase the reliability of assessments.
RoB 2 is based on wide consultation within and outside Cochrane, extensive piloting, and integration of feedback based on user experience. Assessments are made in five bias domains, within which answers to signalling questions address a broader range of issues than in the original RoB tool. These issues include whether post-randomisation deviations from intervention caused bias in trials in which blinding was either not feasible or not implemented and whether outcome data were missing for reasons likely to lead to bias. Assessment of selective reporting is focused on a reported result for an outcome, rather than selective non-reporting of other outcomes that were measured in the trial. RoB 2 also incorporates recent developments in estimation of intervention effects from randomised trials: we distinguish bias in the effect of assignment to interventions from bias in the effect of adhering to intervention as specified in the trial protocol.

RoB 2 assessments relate to the risk of bias in a single estimate of intervention effect for a single outcome or endpoint, rather than for a whole trial. This specificity is because the risk of bias is outcome specific for domains such as bias in measurement of the outcome, and could be specific to a particular estimate (eg, when both intention-to-treat and per protocol analyses have been conducted). We recommend that overall RoB 2 judgments of risk of bias for individual results should be the primary means of distinguishing...
stronger from weaker evidence in the context of a meta-analysis (or other synthesis) of randomised trials. The overall judgments should also influence the strength of conclusions drawn from a systematic review (potentially as part of a GRADE assessment). We strongly encourage stratification by overall risk-of-bias judgment as a default meta-analysis strategy, as shown in figure 2. To facilitate this, we suggest that software for systematic review preparation provides data fields for risk-of-bias assessments. We are preparing an interactive web tool for completing RoB 2 assessments, which we hope will interface well with other systematic review software.

In RoB 2, judgments about risk of bias are derived by algorithms on the basis of answers to specific signalling questions. The added structure provided by the signalling questions aims to make the assessment easier and more efficient to use, as well as to improve agreement between assessors. We believe that this approach is more straightforward than the direct judgments about risk of bias required in the original RoB tool. The algorithms include explicit mappings for situations where there is no information to answer a signalling question, which do not necessarily map to a negative assessment of the trial. For example, when randomisation methods are described and are adequate, the response to the signalling question about baseline imbalances between intervention groups leads to low risk of bias either when such imbalances are compatible with chance, or when there is no information about baseline imbalances. We removed the option of an “unclear” judgment in favour of a graded set of response options (from “low” to “some concerns” to “high”). We envisage that systematic reviews will report the domain level judgments and overall risk-of-bias judgments in tables or figures contained in the main review text. In addition, we encourage reporting of answers to signalling questions, together with direct quotes from papers and free text justification of the answers, in an appendix.

We expect the refinements we have made to the RoB tool to lead to a greater proportion of trial results being assessed as at low risk of bias, because our algorithms map some circumstances to a low risk of bias when users of the previous tool would typically have assessed them to be at unclear (or even high) risk of bias. This potential difference in judgments in RoB 2 compared with the original tool is particularly the case for unblinded trials, where risk of bias in the effect of assignment to intervention due to deviations from intended interventions might be low despite many users of the original RoB tool assigning a high risk of bias in the corresponding domain. We believe that judgments of low risk of bias should be readily achievable for a randomised trial, a study design that is scientifically strong, well understood, and often well implemented in practice. We hope that RoB 2 will be useful to systematic review authors and those making use of reviews, by providing a coherent framework for understanding and identifying trials at risk of bias. This framework might also help those designing, conducting, and reporting randomised trials to achieve the most reliable findings possible.

Author affiliations

1Population Health Sciences, Bristol Medical School, University of Bristol, Bristol BS8 2BN, UK
2NIHR Bristol Biomedical Research Centre, Bristol, UK
3NIHR CLAHRC West, University Hospitals Bristol NHS Foundation Trust, Bristol, UK
4School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia
5Methods team, Epidemiology and Biostatistics Centre, INSERM UMR 1153, Paris, France
6Paris Descartes University, Paris, France
7Cochrane France, Paris, France
8Population Health Research Institute, St George’s, University of London, London, UK
9Centre for Reviews and Dissemination, University of York, York, UK
10Pragmatic Clinical Trials Unit, Centre for Primary Care and Public Health, Queen Mary University of London, UK
11MRC Population Health Research Unit, Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK
12Departments of Epidemiology and Biostatistics, Harvard T H Chan School of Public Health, Harvard-MIT Division of Health Sciences of Technology, Boston, MA, USA
13Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK
14Centre for Evidence-Based Medicine Odense, Odense University Hospital, Odense, Denmark
15Department of Clinical Research, University of Southern Denmark, Odense, Denmark
16Open Patient data Explorative Network, Odense University Hospital, Odense, Denmark
17Department of Emergency Medicine, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada
18Applied Health Research Centre, Li Ka Shing Knowledge Institute, St Michael’s Hospital, Department of Medicine and Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
19Centre for Biostatistics, University of Manchester, Manchester, UK
20Editorial and Methods Department, Cochrane Central Executive, London, UK
21Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA
22Translational Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK
23Nuffield Department of Population Health, University of Oxford, Oxford, UK
24Centre for Clinical Epidemiology, Lady Davis Institute, Jewish General Hospital, McGill University, Montreal, Quebec, Canada
25MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
26MRC Clinical Trials Unit, University College London, London, UK

We dedicate this work to Douglas G Altman, whose contributions were of fundamental importance to development of risk-of-bias assessment in systematic reviews. We thank Henning Keinke Andersen, Nancy Berkman, Mike Campbell, Rachel Churchill, Mike Clarke, Nicky Cullum, François Curtis, Amy Drahotha, Bruno Giraudet, Jeremy Grimshaw, Shaera Ijaz, Yoon Loke, Geraldine Maconald, Richard Morris, Mona Nasser, Nishith Patel, Jani Ruotsalainen, Stephen Senn, Holger Schünemann, Nandi Siegfried, Jayne Tierney, and Sunita Vohra for contributing to discussions; and Andrew Beswick, Julia Bidonde, Angela Busch, staff at Cochrane Argentina, Karen Dawe, Franco De Crescenzo, Kristine Egberts, Clovis Mariano Faggion Jr, Clare French, Lina Golz, Valerie Hoffman, Jon Jackson, Tim Jones, Kayleigh Kew, Elsa Marques, Silvia Minozzi, Theresa Moore, Rebecca Normansell, Rosanne Freak-Poli, Sarah Lensen, José López-López, Marlies Manders, Luke McGuinness, Spyros Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera Papageorgiou, Melissa Randall, Phil Riley, Claudia Smeets, Meera
Provenance and peer review: Not commissioned, externally peer reviewed.

Patient and public involvement: Patients and the public were not involved in this methodological research. We plan to disseminate the research widely, including to community participants in Cochrane.