Introducing a revised tool for assessing risk of bias in randomized trials (RoB 2.0)

Jelena Savović
School of Social and Community Medicine, University of Bristol, UK

With special thanks to Julian Higgins, Matt Page, Jonathan Sterne, Roy Elbers, Barney Reeves, Asbjørn Hróbjartsson, Isabelle Boutron and all RoB 2.0 collaborators
Assessing risk of bias in included studies

Edited by Julian PT Higgins and Douglas G Altman on behalf of the Cochrane Statistical Methods Group and the Cochrane Bias Methods Group

Key Points

- Problems with the design and execution of individual studies of healthcare interventions raise questions about the validity of their findings; empirical evidence provides support for this concern.

- An assessment of the validity of studies included in a Cochrane review should emphasize the risk of bias in their results, i.e. the risk that they will overestimate or underestimate the true intervention effect.

- Numerous tools are available for assessing methodological quality of clinical trials. We recommend against the use of scales yielding a summary score.

- The Cochrane Collaboration recommends a specific tool for assessing risk of bias in each included study. This comprises a description and a judgement for each entry in a ‘Risk of bias’ table, where each entry addresses a specific feature of the study. The judgement for each entry involves answering a question, with answers ‘Yes’ indicating low risk of bias, ‘No’ indicating high risk of bias, and ‘Unclear’ indicating either lack of information or uncertainty over the potential for bias.

SUMMARY POINTS

Systematic reviews should carefully consider the potential limitations of the included studies included.

The Cochrane Collaboration has developed a new tool for assessing risk of bias in randomised trials.

The tool requires a judgment about risk of bias from a description of the study for that judgment, for a series of items covering different domains of bias.
<table>
<thead>
<tr>
<th>Bias</th>
<th>Authors’ judgement</th>
<th>Support for judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation (selection bias)</td>
<td>Unclear risk</td>
<td>Quote: “Subjects were randomised in blocks of six to one of the two treatment groups using sequentially numbered, sealed, opaque envelopes.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comment: sequence generation not reported.</td>
</tr>
<tr>
<td>Allocation concealment (selection bias)</td>
<td>Low risk</td>
<td>Quote: “Subjects were randomised in blocks of six to one of the two treatment groups using sequentially numbered, sealed, opaque envelopes.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comment: allocation process adequate.</td>
</tr>
<tr>
<td>Blinding of participants and personnel (performance bias)</td>
<td>High risk</td>
<td>Quote: “Because the study was not blinded, secondary absorbent dressing and peri ulcer treatments used were at the discretion of the investigator.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comment: stated as not being blinded.</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment (detection bias)</td>
<td>High risk</td>
<td>Quote: “Because the study was not blinded, secondary absorbent dressing and peri ulcer treatments used were at the discretion of the investigator.”</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comment: stated as not being blinded.</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incomplete outcome data (attrition bias)</td>
<td>High risk</td>
<td>Comment: numbers withdrawing and reasons reported by group (Group 1: 14/60 (23%); Group 2: 5/58 (9%)) but a higher proportion of participants withdrew from Group 2 and analysis not undertaken as ITT.</td>
</tr>
<tr>
<td>All outcomes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Selective reporting (reporting bias)</td>
<td>Unclear risk</td>
<td>Comment: although all trial outcomes described in the published report are in the supplied RCT protocol, it was unclear from the published report what the primary outcomes were (maceration in the protocol). A secondary outcome of ‘ability to adapt’ in the protocol (translated from Danish) is not identifiable in the published report.</td>
</tr>
</tbody>
</table>
Current Cochrane tool for risk of bias in randomized trials

- Cochrane RoB tool is very widely used (Jørgensen 2016)
  - 100 out of 100 Cochrane reviews from 2014 (100%)
  - 31 out of 81 non-Cochrane review (38%)
  - >2700 citations from non-Cochrane sources

- The scientific debate on risk of bias has continued

- Evaluation studies of the tool
  - User experience: survey and focus groups (Savovic 2014)
  - Inter-agreement studies (e.g. Hartling 2009 & 2013)
  - Actual use in reviews and published comments (Jørgensen 2016)
Some issues raised with existing tool

- Used *simplistically*
- Used *inconsistently* (domains added or removed)
- Modest *agreement* rates
- RoB judgements are *difficult* for some domains
- Challenges with *unblinded trials*
- Not well suited to *cross-over trials or cluster-randomized trials*
- Not well set up to assess *overall risk of bias*
The revised tool for randomized trials (RoB 2.0) is supported by the UK Medical Research Council Network of Hubs for Trials Methodology Research (MR/L004933/1-N61)
• Revision of the RoB tool started in May 2015
• 1\textsuperscript{st} Development meeting held in Bristol in August 2015
• First ‘working draft’ of the tool completed January 2016
• Piloting phase Feb – March 2016
• Revised ‘working draft’
• 2\textsuperscript{nd} Development meeting to be held in Bristol on 21-22 April 2016
• Development of further guidance and piloting
• Launch at the Seoul Colloquium
RoB 2.0: contributors

- **Core group:**
  - Jelena Savović, Julian Higgins, Matthew Page, Asbjørn Hróbjartsson, Isabelle Boutron, Barney Reeves, Roy Elbers, Jonathan Sterne

- **Working Group members:**
  - Doug Altman, Natalie Blencowe, Mike Campbell, Christopher Cates, Rachel Churchill, Mark Corbett, Nicky Cullum, Francois Curtin, Amy Drahota, Sandra Eldridge, Jonathan Emberson, Bruno Giraudeau, Jeremy Grimshaw, Sharea Ijaz, Sally Hopewell, Asbjørn Hróbjartsson, Peter Jüni, Jamie Kirkham, Toby Lasserson, Tianjing Li, Stephen Senn, Sasha Shepperd, Ian Shrier, Nandi Siegfried, Lesley Stewart, Penny Whiting

- And: Henning Keinke Andersen, Mike Clarke, Jon Deeks, Geraldine MacDonald, Richard Morris, Mona Nasser, Nishith Patel, Jani Ruotsalainen, Holger Schünemann, Jayne Tierney
<table>
<thead>
<tr>
<th>RoB 1.0</th>
<th>RoB 2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random sequence generation</td>
<td>Bias arising from the randomization</td>
</tr>
<tr>
<td>(selection bias)</td>
<td>process</td>
</tr>
<tr>
<td>Allocation concealment</td>
<td>Bias due to deviations from intended</td>
</tr>
<tr>
<td>(selection bias)</td>
<td>interventions</td>
</tr>
<tr>
<td>Blinding of participants and personnel</td>
<td>Bias due to missing outcome data</td>
</tr>
<tr>
<td>(performance bias)</td>
<td></td>
</tr>
<tr>
<td>Incomplete outcome data</td>
<td></td>
</tr>
<tr>
<td>(attrition bias)</td>
<td></td>
</tr>
<tr>
<td>Blinding of outcome assessment</td>
<td>Bias in measurement of the outcome</td>
</tr>
<tr>
<td>(detection bias)</td>
<td></td>
</tr>
<tr>
<td>Selective reporting</td>
<td>Bias in selection of the reported result</td>
</tr>
<tr>
<td>(reporting bias)</td>
<td></td>
</tr>
<tr>
<td>Other bias</td>
<td>N/A</td>
</tr>
<tr>
<td>N/A</td>
<td>Overall bias</td>
</tr>
</tbody>
</table>
Proposed domains of assessment

• All domains to be mandatory

• No additional domains to be available (i.e. no ‘Other bias’ domain)
  • The domains in the tool should cover all potential issues.

• Funding and vested interests to be addressed but not to contribute to overall risk of bias assessments
  • working group led by Asbjørn Hróbjartsson and Isabelle Boutron
Signalling questions and judgements

• Signalling questions are introduced to make the tool easier (and more transparent)
  • ‘Yes’, ‘Probably yes’, ‘Probably no’, ‘No’, ‘No information’

• Risk of bias judgements follow from answers to signalling questions (can be over-ridden)
  • ‘Low risk of bias’, ‘Some concerns’, ‘High risk of bias’

• A change in the interpretation of the judgements, so that a ‘High risk of bias’ judgement in one domain puts the whole study at high risk of bias

• Overall risk of bias judgement can then be completed automatically (can be over-ridden)
<table>
<thead>
<tr>
<th>Bias arising from the randomization process</th>
<th>1.1 Was the allocation sequence random?</th>
<th>Y / PY / PN / N / NI [Description]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?</td>
<td>Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>1.3 Were there baseline imbalances that suggest a problem with the randomization process?</td>
<td>Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>Risk of bias judgement</td>
<td>Low / High / Some concerns [Support]</td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias arising from the randomization process?</td>
<td>[Rationale]</td>
<td></td>
</tr>
<tr>
<td>Bias due to deviations from intended interventions</td>
<td>2.1. Were participants aware of their assigned intervention during the trial?</td>
<td>Y / PY / PN / N / NI [Description]</td>
</tr>
<tr>
<td>2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial?</td>
<td>Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>2.3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice?</td>
<td>NA / Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>2.4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?</td>
<td>NA / Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>2.5 Were any participants analysed in a group different from the one to which they were assigned?</td>
<td>Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>2.6 If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group?</td>
<td>NA / Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>Risk of bias judgement</td>
<td>Low / High / Some concerns [Support]</td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias due to deviations from intended interventions?</td>
<td>[Rationale]</td>
<td></td>
</tr>
<tr>
<td>Bias due to missing outcome data</td>
<td>3.1 Were outcome data available for all, or nearly all, participants randomized?</td>
<td>Y / PY / PN / N / NI [Description]</td>
</tr>
<tr>
<td>3.2. If N/PN/NI to 3.1: Are the proportions of missing outcome data and reasons for missing outcome data similar across intervention groups?</td>
<td>NA / Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>3.3 IfN/PN/NI to 3.2: Is there evidence that results were robust to the presence of missing outcome data?</td>
<td>NA / Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>Risk of bias judgement</td>
<td>Low / High / Some concerns [Support]</td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias due to missing outcome data?</td>
<td>[Rationale]</td>
<td></td>
</tr>
<tr>
<td>Bias in measurement of the outcome</td>
<td>4.1 Were outcome assessors aware of the intervention received by study participants?</td>
<td>Y / PY / PN / N / NI [Description]</td>
</tr>
<tr>
<td>4.2 If Y/PY/NI to 4.1: Was the assessment of the outcome likely to be influenced by knowledge of intervention received?</td>
<td>NA / Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>Risk of bias judgement</td>
<td>Low / High / Some concerns [Support]</td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias due to measurement of the outcome?</td>
<td>[Rationale]</td>
<td></td>
</tr>
<tr>
<td>Bias in selection of the reported result</td>
<td>Are the reported outcome data likely to have been selected, on the basis of the results, from...</td>
<td>Y / PY / PN / N / NI [Description]</td>
</tr>
<tr>
<td>5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?</td>
<td>Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>5.2 ... multiple analyses of the data?</td>
<td>Y / PY / PN / N / NI [Description]</td>
<td></td>
</tr>
<tr>
<td>Risk of bias judgement</td>
<td>Low / High / Some concerns [Support]</td>
<td></td>
</tr>
<tr>
<td>Optional: What is the predicted direction of bias due to selection of the reported result?</td>
<td>[Rationale]</td>
<td></td>
</tr>
<tr>
<td>Overall bias</td>
<td>Risk of bias judgement</td>
<td>Low / High / Some concerns [Support]</td>
</tr>
<tr>
<td>Optional: What is the overall predicted direction of bias for this outcome?</td>
<td>[Rationale]</td>
<td></td>
</tr>
<tr>
<td>Overall risk of bias judgement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Low risk of bias</strong></td>
<td>The study is judged to be at <strong>low risk of bias for all domains</strong> for this result.</td>
<td></td>
</tr>
<tr>
<td><strong>Some concerns</strong></td>
<td>The study is judged to be at <strong>some concerns</strong> in at least one domain for this result.</td>
<td></td>
</tr>
</tbody>
</table>
| **High risk of bias** | The study is judged to be at **high risk of bias** in at least one domain for this result.  
OR  
The study is judged to have **some concerns** for **multiple domains** in a way that substantially lowers confidence in the result. |
Some excerpts from the tool
The RoB 2.0 tool (individually randomized, parallel group trials)

Study design

☑ Randomized parallel group trial
☐ Cluster-randomized trial
☐ Randomized cross-over or other matched design

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 your aim for this study...?

☐ to assess the effect of assignment to intervention
☐ to assess the effect of starting and adhering to intervention
Which of the following sources have you obtained to help inform your risk of bias judgements (tick as many as apply)?

- Journal article(s) with results of the trial
- Trial protocol
- Statistical analysis plan (SAP)
- Non-commercial trial registry record (e.g. ClinicalTrials.gov record)
- Company-owned trial registry record (e.g. GSK Clinical Study Register record)
- “Grey literature” (e.g. unpublished thesis)
- Conference abstract(s) about the trial
- Regulatory document (e.g. Clinical Study Report, Drug Approval Package)
- Research ethics application
- Grant database summary (e.g. NIH RePORTER, Research Councils UK Gateway to Research)
- Personal communication with trialist
- Personal communication with the sponsor
1.1 Was the allocation sequence random?
1.2 Was the allocation sequence concealed until participants were recruited and assigned to interventions?
1.3 Were there baseline imbalances that suggest a problem with the randomization process?
1.2 Was the allocation sequence concealed?

1.1 Was the allocation sequence random?

1.3 Were there baseline imbalances that suggest a problem with randomization?

- Y/PY
- N/PN

NI

Low risk

Some concerns

Some concerns*

High risk

Any response

High risk
## Effect of assignment to intervention

1. Were participants aware of their assigned intervention during the trial?  
2. Were carers and trial personnel aware of participants' assigned intervention during the trial?  
3. If Y/PY/NI to 2.1 or 2.2: Were there deviations from the intended intervention beyond what would be expected in usual practice?  
4. If Y/PY to 2.3: Were these deviations from intended intervention unbalanced between groups and likely to have affected the outcome?  
5. Were any participants analysed in a group different from the one to which they were assigned?  
6. If Y/PY/NI to 2.5: Was there potential for a substantial impact (on the estimated effect of intervention) of analysing participants in the wrong group?
### Effect of starting and adhering to intervention

<table>
<thead>
<tr>
<th>Question</th>
<th>Blinding</th>
<th>Specific deviations</th>
<th>Overcome by analysis?</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Were participants aware of their assigned intervention during the trial?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.2. Were carers and trial personnel aware of participants' assigned intervention during the trial?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.3. If Y/PY/NI to 2.1 or 2.2: Were important co-interventions balanced across intervention groups?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.4. Was the intervention implemented successfully?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5. Did study participants adhere to the assigned intervention regimen?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.6. If N/PN/NI to 2.3, 2.4 or 2.5: Was an appropriate analysis used to estimate the effect of starting and adhering to the intervention?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Bias in selection of the reported result

Are the reported outcome data likely to have been selected, on the basis of the results, from...

5.1. ... multiple outcome measurements (e.g. scales, definitions, time points) within the outcome domain?

5.2 ... multiple analyses of the data?

Selective outcome reporting
Selective analysis reporting
RoB 2.0 tool

A revised tool to assess risk of bias in randomized trials (RoB 2.0)

Welcome to the website for the RoB 2.0 tool. This is a draft version of the tool. We have developed versions for three different trial designs.

Individually randomized, parallel group trials

You can:

- Download background information and detailed guidance for using the RoB 2.0 tool (pdf).
- Download the tool itself (pdf).
- Download a blank template for completing the tool, which has two variants
  - Implement RoB 2.0 when interest is in the effect of assignment to intervention (Word).
  - Implement RoB 2.0 when the interest is in the effect of starting and adhering to intervention (Word).

Cluster randomized, parallel group trials
The effect of interest

- The current tool has nothing to say about review questions for which blinding is not feasible
  - (other than to classify as not blind hence high risk of bias)
- Issues of performance bias very different for "ITT effects" and "per-protocol" effects, yet poorly addressed in current RoB tool
- "ITT effect": effect of assignment to intervention
  - e.g. the question of interest to a policy maker about whether to introduce a screening programme
- "Per protocol effect": effect of starting and adhering to intervention
  - e.g. the question of interest to an individual about whether to attend screening

Not to be confused with ITT or per protocol analyses
Effect of interest

• Deviations from intended intervention are not important when interest is on the effect of assignment to intervention
  • e.g. some people don’t respond to invitations to be screened
  • ...providing these deviations reflect routine care
    • rather than behaviour that reflects expectations of a difference between intervention and comparator

• But deviations such as poor adherence, poor implementation and co-interventions may lead to bias when interest is in the effect starting and adhering to intervention

• We therefore have different tools for these two effects of interest
Bias in selection of the reported result
Outcome non-reporting bias

- Current tool emphasises assessment of selective non-reporting or partial reporting of outcomes:
  - e.g. trialists measure pain, function and QoL, but only report data for pain
  - e.g. trialists report P values but no means & SDs for pain

- Review authors often rate a study at high risk of bias if one outcome is not reported
  - e.g. “All outcomes were reported except for pain”
  - e.g. “Some outcomes were not reported”
2 trials are rated at high risk of bias because pain was not reported

But this is a meta-analysis of function, so it does not make sense to display these high risk ratings here
Selective non-reporting biases the result of the **meta-analysis** which cannot include the trial that omitted the outcome; it does not bias the trial result.

This is similar to publication bias (non-reporting of a study).
Bias in selection of the reported result

Trial result is biased because it has been selected on the basis of the results from multiple:

- Outcome measurements
  - Scales
  - Definitions of/criteria for an event
  - Time points
- Analyses
  - Unadjusted vs adjusted models
  - Different sets of covariates in adjusted models
  - Final values vs change from baseline vs analysis of covariance
  - Continuous scale converted to categorical data with different cut-points
We propose that:

- Bias in selection of the reported result be addressed in the revised risk of bias tool

- Selective non-reporting (and partial reporting) of outcomes be addressed elsewhere, in a new tool to assess the risk of reporting biases in meta-analyses
Cluster-randomized trials and cross-over trials

- Cluster-randomized trials:
  - Key issue is recruitment / identification of participants after interventions have been allocated to clusters
  - Also consideration of missing data at cluster and individual level

- Cross-over trials (AB/BA design)
  - Key issue is carry-over of effect from 1\textsuperscript{st} period to 2\textsuperscript{nd} period
  - Also period effects, selective reporting of 1\textsuperscript{st} period data
Some unresolved issues

• How many results to assess per study?

• How much free text to include to support assessments?

• How should it be presented in the review?

• Implementation
  • RoB 2.0 will need careful consideration to make the process efficient for multiple outcomes
  • Discussions to be initiated with RevMan team at Seoul Colloquium
• We believe RoB 2.0 offers considerable advantages over the existing tool
• Once programmed into software, we expect the tool will be easier to use than the first one
• We are extremely grateful to all those who have contributed to the development of RoB 2.0
• RoB 2.0 is available at www.riskofbias.info