RoB 2 Domain 5: Bias in selection of the reported result

Matthew Page and Isabelle Boutron
Monash University, Australia
Université de Paris, France

Trusted evidence.
Informed decisions.
Better health.
• Selective reporting vs selective non-reporting
• Types of bias in selection of the reported result
• Questions
• How to spot bias in selection of the reported result?
• Questions
• Assessing the risk of bias in selection of the reported result in RoB 2
• Examples
• Questions and discussion
Risk of bias in randomized trials

- Bias arising from the randomization process
- Bias due to deviations from intended intervention
- Bias due to missing outcome data
- Bias in measurement of the outcome
- Bias in selection of the reported result

Experimental
Comparator

Outcome

1.02  3.87
2.20  4.32
1.38  5.44
Risk of bias in randomized trials

- Bias arising from the randomization process
- Bias due to deviations from intended intervention
- Bias due to missing outcome data
- Bias in measurement of the outcome

Experimental
Comparator
Outcome
Outcome

1.02 3.87
2.20 4.32
1.38 5.44

Bias in selection of the reported result
• Have you ever used the original Cochrane risk of bias tool for randomized trials?
Selective reporting vs selective non-reporting

1 result reported from among 3 scales used for the same clinical outcome

Potential bias in selection of the reported result: Domain 5 in RoB 2

Not reported, because “P > 0.05”

Potential bias in the meta-analysis: ROB-ME tool

Study 1
Study 2
Study 3
Study 4
Study 5
Study 6
Selective non-reporting addressed in the ROB-ME (“Risk Of Bias due to Missing Evidence”) tool

ROB-ME integrates assessment of risk of bias in *meta-analyses* due to:
- Missing studies (publication bias)
- Missing results (selective non-reporting bias)

beta version of ROB-ME tool to be launched October 27 at [riskofbias.info](http://riskofbias.info)
Bias in selection of the reported result

- **Outcome domain**
  - **Outcome measurement**
    - **Outcome analysis**

Examples:
- e.g. severity of depression
- e.g. Hamilton rating scale after 6 weeks
- e.g. difference in mean change in Hamilton score from baseline to 6 weeks
Bias in selection of the reported result

There are multiple possible results that could be generated for a domain, each of which are eligible for the synthesis we want to undertake.

e.g. severity of depression

e.g. Hamilton rating scale after 6 weeks

e.g. difference in mean change in Hamilton score from baseline to 6 weeks
Bias in selection of the reported result

Bias may arise when results are selected based on their magnitude, direction or P value, from:

- multiple outcome *measurements* within the outcome domain, e.g.
  - multiple scales
  - multiple definitions of/criteria for an event
  - multiple time points
Bias in selection of the reported result

Bias may arise when results are selected based on their magnitude, direction or P value, from:

- multiple *analyses* of the outcome measurement, e.g.
  - 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
Questions
How to spot bias in selection of the reported result

• Preliminary considerations

Which of the following sources were obtained to help inform the risk-of-bias assessment? (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 or Research Councils UK Gateway to Research)
☐ Personal communication with trialist
☐ Personal communication with the sponsor
How to spot bias in selection of the reported result

- ideally, we have a pre-specified analysis plan
  - a trial protocol (even a detailed trial registry record) may be sufficient
  - statistical analysis plan (SAP) often provides the most detail

- check for any amendments or updates to plans
  - usually date-stamped in trials registry or journal publication
How to spot bias in selection of the reported result

- analysis plan should be date-stamped, confirming that planned analyses were finalised before unblinded outcome data were made available for analysis

<table>
<thead>
<tr>
<th>Protocol Number Code:</th>
<th>DMID Protocol: 20-0006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Development Phase:</td>
<td>Phase 3</td>
</tr>
</tbody>
</table>
| Products:             | Remdesivir
                      | Placebo                |
| Form/Route:           | IV                     |
| Indication Studied:   | COVID-19               |
| Sponsor:              | Division of Microbiology and Infectious Diseases
                      | National Institute of Allergy and Infectious Diseases
                      | National Institutes of Health |
| Clinical Trial Initiation Date: | February 21, 2020 |
| Clinical Trial Completion Date: | Trial Ongoing |
| Date of the Analysis Plan: | April 20, 2020 |
| Version Number:       | 1.0                    |
How to spot bias in selection of the reported result

- If there is an analysis plan, check the trial report agrees with it:
  - outcome measures changed?
  - analysis methods changed?
  - any explanation for the changes?

- Reminder: focus only on changes relating to the trial result(s) being assessed for risk of bias
How to spot bias in selection of the reported result

- Some differences between analysis plan and trial report may be due to legitimate changes to the protocol:
  - planned cut-points for a continuous outcome needed to be modified because the distribution of data differed to what was anticipated
  - timing of follow-up was delayed because the measurement device was broken
  - plans were modified before conducting any analyses, yet protocol not updated

- Contact trialists for clarification
How to spot bias in selection of the reported result

- what if there is no pre-specified analysis plan?
  - compare ‘Methods’ with ‘Results’ – look for:
    - whether outcome measurements or analyses reported match what was described in the ‘Methods’ section
    - any measurements or analyses added that were apparently not planned
How to spot bias in selection of the reported result

• what if there is no pre-specified analysis plan?
  – consider the following questions:
    • were outcome measures and analyses consistent across multiple reports relating to a study?
    • are subscales aggregated in an unusual manner?
    • have the researchers categorized continuous outcome measures in an unusual way?
    • has an unusual combination of unanticipated adverse events been categorised as “serious” and “minor” adverse events?
How to spot bias in selection of the reported result

- Take all sources (e.g. protocol, journal article, clinical study report, results supplied by authors) into consideration when reaching a judgement of risk of bias
- No reason for concern if, across all sources, you have access to results of all planned analyses
Questions
Assessing the risk of bias in selection of the reported result
Bias in selection of the reported result

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?
Reponses: Y/PY/PN/N/NI

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?
Reponses: Y/PY/PN/N/NI

5.3 ... multiple eligible analyses of the data?
Reponses: Y/PY/PN/N/NI
Is the numerical result being assessed likely to have been selected, on the basis of the results, from...
- multiple eligible outcome measurements
- multiple analysis

‘No’ or ‘Probably no’:

• All reported results are consistent with what was planned.
• Only one possible way in which the outcome domain can be measured / analyzed
• All inconsistencies are explained and not related to the results.

‘Yes’ or ‘Probably yes’:

• Clear evidence that the results reported were selected on the basis of the results (e.g., statistically / non statistically significant).
Bias in selection of the reported result

<table>
<thead>
<tr>
<th>5.1 (In accordance with plan?)</th>
<th>5.2 (Selected from multiple outcomes?)</th>
<th>5.3 (Selected from multiple analyses?)</th>
<th>Domain level judgement</th>
<th>Default risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y/PY</td>
<td>N/PN</td>
<td>N/PN</td>
<td></td>
<td>Low</td>
</tr>
<tr>
<td>N/PN/NI</td>
<td>N/PN</td>
<td>N/PN</td>
<td></td>
<td>Some concerns</td>
</tr>
<tr>
<td>Any answer</td>
<td>N/PN</td>
<td>NI</td>
<td></td>
<td>Some concerns</td>
</tr>
<tr>
<td>Any answer</td>
<td>NI</td>
<td>N/PN</td>
<td></td>
<td>Some concerns</td>
</tr>
<tr>
<td>Any answer</td>
<td>NI</td>
<td>NI</td>
<td></td>
<td>Some concerns</td>
</tr>
<tr>
<td>Any answer</td>
<td>Either 5.2 or 5.3 Y/PY</td>
<td></td>
<td></td>
<td>High</td>
</tr>
</tbody>
</table>

Y/PY = ‘Yes’ or ‘Probably yes’; N/PN = ‘No’ or ‘Probably no’; NI = ‘No information’
Bias in selection of the reported result

5.1 Trial analysed in accordance with a pre-specified plan?

- Y/PY → Low risk
- N/PN/NI → Some concerns

Result selected from...

5.2 ...multiple outcome measurements? [Both N/PN]

- Y/PY → Low risk
- N/PN → Some concerns

5.3 ...multiple analyses of the data?

- Either Y/PY → High risk
### Selection of the reported result

#### Signalling

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?
- **Response**: NI

#### Risk of bias judgement

<table>
<thead>
<tr>
<th>Algorithm result</th>
<th>Assessor's judgement</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>

Optional: What is the predicted direction of bias due to selection of the reported result?
- **Response**: 

**Description**

- **Response**: 

---

### RoB 2 assessment for individual randomized, parallel group trials

- **Unique ID (e.g. A1 or 1)**
- **Study ID**
- **Experimantal Comparator**
- **Specify which outcome**
- **Specify the numerical result**

**Weight for analysis**

- **Is the review team's aim for this results to assess...?**
- **If the aim is to assess the effect of adhering to intervention... (select one at least)**
  - Occurrence of non-protocol interventions
  - Failure in implementing the intervention that could have affected the outcome

**Domain 1 | Domain 2 | Domain 3 | Domain 4 | Domain 5 | Overall bias**

---

**Which of the following sources were obtained to help inform the risk-of-bias assessment? (tick as many as apply; for editing, please double-click the list)**
- 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
Bias in selection of the reported result

Low risk of bias

• prespecified trial analysis plan is available
• eligible outcome measures (and analyses) reported according to trial analysis plan, irrespective of the results

High risk of bias

• evidence (or strong hint) that the reported outcome measure (or analysis) was selected on the basis of the results

Some concerns

• many trials will be judged in this category
Randomized controlled trial evaluating intradiscal injection of steroid in patients with low back pain

*Figure 2. Mean lumbar pain intensity in previous 48 h, by intervention group.*
Randomized controlled trial evaluating intradiscal injection of steroid in patients with low back pain

Figure 2. Mean lumbar pain intensity in previous 48 h, by intervention group.

Different possible ways to report the results?
Reporting of results?

- Mean at M1
- Mean at M6
- Mean at M12

- Mean change from baseline at M1
- Mean change from baseline at M6
- Mean change from baseline at M12

Dichotomization

- Success is defined as less than 40/100 on pain numeric scale
- Success is defined as less than 35/100 on pain numeric scale
- Success is defined as less than 30/100 on pain numeric scale
- Etc…
Outcome in the registry/Protocol/SAP (blinded and before the analysis)

**Primary Outcome Measures**: Back pain level assessed on a 11-point numeric scale (0-100) at 1 month. Success is defined as less than 40 on pain numeric scale at 1 month [ Time Frame: 1 month ]

CT.gov

---

Outcome in the publication

The primary outcome was the percentage of patients with LBP intensity less than 40 on an 11-point numerical rating scale (0 [no pain] to 100 [maximum pain] in 10-point increments) in the previous 48 hours at 1 month after the intervention. The main secondary outcomes were LBP intensity and
Outcome in the registry/Protocol/SAP (blinded and before the analysis)

**Primary Outcome Measures**: Back pain level assessed on a 11-point numeric scale (0-100) at 1 month. Success is defined as less than 40 on pain numeric scale at 1 month [Time Frame: 1 month]

CT.gov

Outcome in the publication

The primary outcome was the percentage of patients with LBP intensity less than 40 on an 11-point numerical rating scale (0 [no pain] to 100 [maximum pain] in 10-point increments) in the previous 48 hours at 1 month after the intervention. The main secondary outcomes were LBP intensity and

Low risk of bias

Analysed in accordance with a prespecified plan
Outcome in the registry/Protocol/SAP

Primary outcome: Back pain level assessed on a 11-point numeric scale (0-100) at 12 month

Outcome in the publication

The primary outcome was the percentage of patients with LBP intensity less than 40 on an 11-point numerical rating scale (0 [no pain] to 100 [maximum pain] in 10-point increments) in the previous 48 hours at 1 month after the intervention. The main secondary outcomes were LBP intensity and
Randomized controlled trial evaluating intradiscal injection of steroid in patients with low back pain

Figure 2. Mean lumbar pain intensity in previous 48 h, by intervention group.
Outcomes in the registry/Protocol/SAP

Primary outcome: Back pain level assessed on a 11-point numeric scale (0-100) at 12 month

Outcome in the publication

The primary outcome was the percentage of patients with LBP intensity less than 40 on an 11-point numerical rating scale (0 [no pain] to 100 [maximum pain] in 10-point increments) in the previous 48 hours at 1 month after the intervention. The main secondary outcomes were LBP intensity and

High risk of bias

The numerical result being assessed likely to have been selected on the basis of the results
COVID example
Horby P, Lancet, 2020

Outcomes
• Mortality (day 28)
• Time to discharge from hospital

Resources
  o Protocol – Yes
  o Statistical Analysis Plan - Yes
  o Registry entry (prospective registration)- Yes
  o All resources are consistent

2.6 Definitions of primary and secondary outcomes
Outcomes will be assessed at 28 days and then 6 months after randomisation. Analysis of longer-term outcomes collected beyond this will be described in a separate Statistical Analysis Plan.

2.6.1 Primary outcome
Mortality (all-cause)

2.6.2 Secondary clinical outcomes
• Time to discharge from hospital
• Use of mechanical ventilation/Extra Corporal Membrane Oxygenation (ECMO) or death (among patients not on ventilation or ECMO at baseline)

2.6.3 Subsidiary clinical outcomes
• Cause-specific mortality (COVID-19; cardiovascular; non-vascular; other)
• Use of renal dialysis or haemofiltration
• Serious cardiac arrhythmia (recorded in a subset)
Horby P, Lancet, 2020

Outcomes
- Mortality (day 28)
- Time to discharge from hospital

Resources
- Protocol – Yes
- Statistical Analysis Plan - Yes
- Registry entry (prospective registration)-Yes
- All resources are consistent

Low risk of bias
Analysed in accordance with a prespecified plan

2.6 Definitions of primary and secondary outcomes
Outcomes will be assessed at 28 days and then 6 months after randomisation. Analysis of longer-term outcomes collected beyond this will be described in a separate Statistical Analysis Plan.

2.6.1 Primary outcome
- Mortality (all-cause)

2.6.2 Secondary clinical outcomes
- Time to discharge from hospital
- Use of mechanical ventilation/Extra Corpuscular Membrane Oxygenation (ECMO) or death (among patients not on ventilation or ECMO at baseline)

2.6.3 Subsidiary clinical outcomes
- Cause-specific mortality (COVID-19; cardiovascular; non-vascular; other)
- Use of renal dialysis or haemofiltration
- Serious cardiac arrhythmia (recorded in a subset)

RECOVERY SAP
Version number: 1.0

- Use of ventilation (overall and by type)
- Duration of ventilation (overall and by type)
Outcomes

• Time to clinical improvement
• Time to viral negative conversion
• Incidence of clinical improvement (day 14)

Sources

• Protocol – no
• Statistical Analysis Plan - no
• Registry entry – Yes

Analysis planned /reported

○ All outcomes are reported completely in the registry and consistently in the publication
COVID example

Outcomes
  • Time to clinical improvement
  • Time to viral negative conversion
  • Incidence of clinical improvement (day 14)

Sources
  • Protocol – no
  • Statistical Analysis Plan - no
  • Registry entry – Yes

Analysis planned /reported
  o All outcomes are reported completely in the registry and consistently in the publication
  o However the registration was not done prospectively

Some concerns.
  NI on the prespecified plan
Questions and discussion