How Cochrane’s decision to include non-randomized studies has led to important methods research

Barney Reeves
Methods research setting
Cochrane’s position about NRSi

- Cochrane has always recognised that NRSi can contribute important information [Oxman et al., 1994]

- “We (Cochrane) gather and summarize the best evidence from research to help you make informed choices about treatment.” [www.cochrane.org/about-us]

- NRSiMG recommendation: Review authors should formally consider whether NRSi are necessary to answer the review question. [Reeves et al., J Res Methods Synth, 2013]

- Recommendation not based on “methods research”, but:
  - many important questions are not addressed by RCTs
  - e.g. in 2012, specific harms outcomes were reported in only 38% of new Cochrane reviews [Saini et al. BMJ 2014]
What’s different when including NRSi?

• Title
• Protocol
• Design of searches / searching
• ‘Triage’ abstracts for eligibility
• ‘Triage’ full papers for eligibility
• Data extraction, including risk of bias (RoB) assessment
• Data synthesis
• Interpretation
Protocol

• Review question
  – What would a RCT of the review question look like ("target trial")? [Sterne et al. www.riskofbias.info]
  – What is the nature of the target comparison? [Sterne et al. www.riskofbias.info]
  – Confounding domains [Sterne et al. www.riskofbias.info]

• Criteria for study eligibility
  – Specify study design features cf. labels [Handbook, Ch.13]

• Plan for synthesis [Handbook, Ch.13]
  – Meta-analyse or not? Forest plots without pooled estimates
  – Adjusted vs unadjusted effect estimates
  – Multiple adjusted effect estimates
Study design features, not labels

Table 1. Checklist of study design features for studies formed by classifying individuals by intervention and comparator.

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Can’t tell</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>Was there a relevant comparison:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Between two or more groups of participants receiving different interventions?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Within the same group of participants over time?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. <strong>Were groups formed by:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Randomization?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quasi-randomization?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other action of researchers?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Time differences?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Location differences?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Healthcare decision makers?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participant preferences?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>On the basis of outcome?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some other process? (specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. <strong>Were the key steps of the study described below carried out after the study was designed:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identification of participants?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessment before intervention?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Actions/choices leading to an individual becoming a member of a group?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessment of outcomes?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. <strong>On which variables was comparability between groups assessed:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Potential confounders?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Assessment of outcome variables before intervention?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
When to include NRSi

- Few if any RCTs will address the PICO? No
- RCTs populations are restrictive so unlikely to address the PICO? No
- The intervention in usual care will differ in important ways so that RCTs are unlikely to address the PICO? No
- The comparator in usual care will differ in important ways so that RCTs are unlikely to address the PICO? No

- ONLY INCLUDE RCTS

- Do NRS use study design features as specified in eligibility criteria? No
- Do NRS address the PICO (especially the intervention)? No
- Do NRS have a “confirmatory” objective (e.g. define a clinically important target difference, etc.)? No
- In NRS, is the basis for “indication” sufficiently understood to judge the risk of confounding? No

INCLUDE NRS [+-RCTS]
Searching for and ‘triaging’ studies

• Searching:
  – Avoid design terms [Handbook, Ch.13]
  – Less comprehensive search?

• ‘Triaging’ abstracts
  – Difficult to exclude abstracts based on abstract

• Final selection from full papers
  – Apply study design checklist [Handbook, Ch.13]
Data extraction

• Assess risk of bias (ACROBAT-NRSi) [Sterne et al. www.riskofbias.info]
  – Study level information: target trial same as for review question? nature of comparison? specific effect to be appraised?
  – Outcome level information: signalling questions, domain-level RoB, outcome-level RoB
Bias due to confounding

1.1 Is confounding of the effect of intervention unlikely in this study?
If Y or PY to 1.1, the study can be considered to be at low risk of bias due to confounding and no further signalling questions need be considered.

1.2. If N or PN to 1.1: Were participants analysed according to their initial intervention group throughout follow-up?
If Y or PY to 1.2, answer questions 1.4 to 1.6, which relate to baseline confounding.

1.3. If N or PN to 1.2: Were intervention discontinuations or switches unlikely to be related to factors that are prognostic for the outcome?
If Y or PY to 1.3, answer questions 1.4 to 1.6, which relate to baseline confounding.

1.4. Did the authors use an appropriate analysis method that adjusted for all the critically important confounding domains?

1.5. If Y or PY to 1.4:
Were confounding domains that were adjusted for measured validly and reliably by the variables available in this study?

1.6. Did the authors avoid adjusting for post-intervention variables?

1.7. Did the authors use an appropriate analysis method that adjusted for all the critically important confounding domains and for time-varying confounding?

1.8. If Y or PY to 1.7:
Were confounding domains that were adjusted for measured validly and reliably by the variables available in this study?

Bias in selection of participants into the study

2.1. Was selection into the study unrelated to intervention or unrelated to outcome?

2.2. Do start of follow-up and start of intervention coincide for most subjects?

2.3. If N or PN to 2.1 or 2.2:
Were adjustment techniques used that are likely to correct for the presence of selection biases?

Bias in measurement of interventions

3.1. Is intervention status well defined?

3.2. Was information on intervention status recorded at the time of intervention?

3.3. Was information on intervention status unaffected by knowledge of the outcome or risk of the outcome?

Risk of bias judgement
(Optional) Predicted direction of bias

Bias due to departures from intended interventions

4.1. Were the critical co-interventions balanced across intervention groups?

4.2. Were numbers of switches to other interventions low?

4.3. Was implementation failure minor?

4.4. If N or PN to 4.1, 4.2 or 4.3:
Were adjustment techniques used that are likely to correct for these concerns?

Risk of bias judgement
(Optional) Predicted direction of bias

Bias due to missing data

5.1. Are outcome data reasonably complete?

5.2. Was intervention status reasonably complete for those in whom it was sought?

5.3. Are data reasonably complete for other variables in the analysis?

5.4. If N or PN to 5.1, 5.2 or 5.3:
Are the proportion of participants and reasons for missing data similar across interventions?

5.5. If N or PN to 5.1, 5.2 or 5.3:
Were appropriate statistical methods used to account for missing data?

Risk of bias judgement
(Optional) Predicted direction of bias

Bias in measurement of outcomes

6.1. Was the outcome measure objective?

6.2. Were outcome assessors unaware of the intervention received by study participants?

6.3. Were the methods of outcome assessment comparable across intervention groups?

6.4. Were any systematic errors in measurement of the outcome unrelated to intervention received?

Risk of bias judgement
(Optional) Predicted direction of bias

Bias in selection of the reported result

Is the reported effect estimate unlikely to be selected, on the basis of the results, from...

7.1 ...among multiple outcome measurements within the outcome domain?

7.2 ...among multiple analyses of the intervention-outcome relationship?

7.3 ...among different subgroups?

Risk of bias judgement
(Optional) Predicted direction of bias

Overall risk of bias

Risk of bias judgement
(Optional) Predicted direction of bias

Sterne et al. www.riskofbias.info
<table>
<thead>
<tr>
<th>Domain</th>
<th>Related terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias due to confounding</td>
<td>Selection bias as it is often used in relation to clinical trials (and currently in widespread use within The Cochrane Collaboration); Allocation bias; Case-mix bias.</td>
</tr>
<tr>
<td>Bias in selection of participants into the study</td>
<td>Selection bias as it is usually used in relation to observational studies; Inception bias; Lead-time bias; Immortal time bias.</td>
</tr>
<tr>
<td>Bias in classification of interventions</td>
<td>Misclassification bias; Information bias; Recall bias; Measurement bias; Observer bias.</td>
</tr>
<tr>
<td>Bias due to departures from intended interventions</td>
<td>Performance bias; Time-varying confounding bias.</td>
</tr>
<tr>
<td>Bias due to missing data</td>
<td>Attrition bias; Selection bias as it is usually used in relation to observational studies.</td>
</tr>
<tr>
<td>Bias in measurement of outcomes</td>
<td>Detection bias; Information bias, Misclassification bias; Observer bias.</td>
</tr>
<tr>
<td>Bias in selection of the reported result</td>
<td>Outcome reporting bias, Analysis reporting bias.</td>
</tr>
</tbody>
</table>
Agenda for the future

Current initiatives include:

• Adapting the algorithm for deciding when to include NRSi so that it can inform GRADE

• Extending study feature checklist to cover types of studies used by health systems, social care and policy researchers

• Validating ACROBAT-NRSi (in collaboration with the Bias and Statistics MG). [Higgins et al.]

Looking further ahead:

• Explore how treatment effects change with searches of varying comprehensiveness.

• Research the risk of confounding and selection of participants in NRSi in different circumstances.