Incorporating ‘risk of bias’ assessments into meta-analyses

Julian Higgins
Where risk of bias might be addressed in a systematic review

1. Formulation of a clear question and eligibility criteria for studies
2. Search for potentially relevant studies
3. Selection of studies into the review
4. Collection of data
5. Assessment of methodological quality of included studies
6. Synthesis of findings (possibly using meta-analysis)
7. Presentation of data and results
8. Interpretation and drawing conclusions

Threshold for eligibility

Information on methods, conduct and limitations needs to be collected

Formal assessment of included studies

Risks of bias should be integrated into analysis

Interpretation and conclusions need to reflect risk of bias in the evidence
Incorporating assessments into analyses

**Only the studies at low risk of bias:**
Low risk of bias, but may be imprecise (wide confidence interval) because of limited information

**All the studies:**
High precision, but may be seriously biased because of flaws in the conduct of some of the studies.
Comparison of subgroups, and meta-regression

- Explore the relationship between component(s) of study quality (e.g. blinding) and overall result
- Avoid using scales
What is meta-regression?

• A method for investigating possible causes of heterogeneity in a meta-analysis

• Relates study characteristics to their sizes of effect

• Problems we can address using meta-regression:
  – Does the intervention work better if given for longer?
  – Are smaller odds ratios observed in high-risk populations?
  – Is inadequate allocation concealment associated with a larger effect estimate?
  – Is there a relationship between sample size and effect size (e.g. due to publication bias)?
Limitations of meta-regression

• Low power in most Cochrane reviews:
  – unable to detect relationships that are there
  – typically cannot learn reliably about biases within one review or meta-analysis

• Spurious findings
  – high risk of false positive results by chance

• Confounding
  – cannot conclude causality
  – since associations are observational

• Can’t be done in RevMan
Sensitivity analysis

• Are results sensitive to decisions and assumptions made to obtain them

• e.g.
  – Same result in ‘better’ studies?
  – Same result if [unblinded/short/high drop-out] studies excluded?
the major approach to incorporating risk of bias assessments in Cochrane reviews is to **restrict** meta-analyses to studies at low (or lower) risk of bias, or to **stratify** studies according to risk of bias.
Options

• Primary analysis restricted to studies at low (or low and unclear) risk of bias
  – always conduct sensitivity analysis

• Present multiple (stratified) analyses

• Present all studies and provide a narrative discussion of risk of bias
  – be missed by readers
  – does not address impact on results
Primary analysis restricted to studies at low (or low and unclear) risk of bias

• Exclude studies with high risk of bias from the synthesis

• Based on summary judgements or one or more key domains

• All thresholds are arbitrary
  – spectrum from ‘free of bias’ to ‘undoubtedly biased’

• Problematic if few trials

• Perform sensitivity analyses showing how conclusions might be affected if studies at high risk of bias were included
What about ‘Unclear risk of bias’?

• It is recommended that review authors do not combine studies at ‘low’ and ‘unclear’ risk of bias in analyses, unless they provide specific reasons for believing that these studies are likely to have been conducted in a manner that avoided bias. In the rest of this section, we will assume that studies assessed as at low risk of bias will be treated as a separate category.
Present multiple (stratified) analyses

- May produce at least three estimates of the intervention effect
  - high risk of bias
  - low risk of bias
  - all studies
- Present two or more with equal prominence?
- May be confusing for readers
- Stratified *forest plots* present all information transparently
Restrict or stratify?

• Based on
  – context of the review
  – balance between the potential for bias and the loss of precision when studies at high or unclear risk of bias are excluded

• Recall, lack of a statistically significant difference between studies at high and low risk of bias should not be interpreted as implying absence of bias
  – because meta-regression analyses typically have low power
Summary of findings table

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Illustrative comparative risks* (95% CI)</th>
<th>Relative effect (95% CI)</th>
<th>No of Participants (studies)</th>
<th>Quality of the evidence (GRADE)</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Assumed risk</td>
<td>Corresponding risk</td>
<td>3a</td>
<td>3b</td>
<td>4</td>
</tr>
<tr>
<td>Control</td>
<td>Glucosamine versus placebo (adequate allocation concealment)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. List of all important outcomes (desirable and undesirable)
2. A measure of the typical burden of these outcomes on control group
3. Absolute and relative magnitude of effect (if both are appropriate)
4. Numbers of participants in studies addressing these outcomes
5. Rating of quality of evidence for each outcome
6. Space for comments
What is GRADE?

• Judgment about the quality of evidence for each main outcome

• Within the context of a systematic review, GRADE reflects how confident we are that an estimate is close to the truth
# GRADE in a nutshell

<table>
<thead>
<tr>
<th>Grade</th>
<th>Confidence and Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Further research is very unlikely to change our confidence in estimate of effect</td>
</tr>
<tr>
<td>Moderate</td>
<td>Further research likely to have impact on confidence in the estimate of effect and may change the estimate</td>
</tr>
<tr>
<td>Low</td>
<td>Further research is very likely to have an important impact on our confidence in the estimate of effect, and is likely to change the estimate</td>
</tr>
<tr>
<td>Very low</td>
<td>Any estimate of effect is very uncertain</td>
</tr>
</tbody>
</table>
GRADE in a nutshell

High

Randomized trials start here

Moderate

5 reasons to downgrade

Low

Observational studies start here

Very low

3 reasons to upgrade

Concentrate on this

Unusual
What might decrease quality of evidence

1. Study limitations (risk of bias)
2. Indirectness of evidence
3. Inconsistency of results
4. Imprecision
5. Publication bias

- Each category for factors that might decrease quality of evidence has 3 scoring options:
  - No concerns
  - Serious −1 level
  - Very serious −2 levels
## Summary risk of bias by outcome

<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Interpretation</th>
<th>Within a study</th>
<th>Across studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low risk of bias</td>
<td>Plausible bias unlikely to seriously alter the results</td>
<td>Low risk of bias for all key items</td>
<td>Most information is from studies at low risk of bias</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
<td>Plausible bias that raises some doubt about the results</td>
<td>Unclear risk of bias for one or more key items</td>
<td>Most information is from studies at low or unclear risk of bias</td>
</tr>
<tr>
<td>High risk of bias</td>
<td>Plausible bias that seriously weakens confidence in the results</td>
<td>High risk of bias for one or more key items</td>
<td>The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results</td>
</tr>
<tr>
<td>Risk of bias</td>
<td>Across studies</td>
<td>Considerations</td>
<td>GRADE assessment</td>
</tr>
<tr>
<td>-----------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td>Low risk of bias</td>
<td>Most information is from studies at low risk of bias</td>
<td>No apparent limitations</td>
<td>No serious limitations, do not downgrade</td>
</tr>
<tr>
<td>Unclear risk of bias</td>
<td>Most information is from studies at low or unclear risk of bias</td>
<td>Most information is from studies at low risk of bias</td>
<td>No serious limitations, do not downgrade</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Potential limitations are unlikely to lower confidence in the estimate of effect.</td>
<td>Serious limitations, downgrade one level</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Potential limitations are likely to lower confidence in the estimate of effect</td>
<td>Serious limitations, downgrade one level</td>
</tr>
<tr>
<td>High risk of bias</td>
<td>The proportion of information from studies at high risk of bias is sufficient to affect the interpretation of results</td>
<td>Crucial limitation for one criterion, or some limitations for multiple criteria, sufficient to lower confidence in the estimate of effect.</td>
<td>Serious limitations, downgrade one level</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crucial limitation for one or more criteria sufficient to substantially lower confidence in the estimate of effect.</td>
<td>Very serious limitations, downgrade two levels</td>
</tr>
</tbody>
</table>
Weighting by quality scores

- Weights based purely on quality scores are arbitrary, have little statistical justification and are *not recommended*
Adjusting for bias

• Bayesian analysis
  – Incorporate prior distributions on likely bias for each individual study

• Prior based on
  – empirical research (e.g. meta-epidemiology): Welton et al
  – plausible ranges
  – formal elicitation from experts: Turner et al

• These are perhaps the ‘state of the art’, but not likely to be feasible for most review teams
  – Save for the most important reviews
Concluding remarks

• Restrict or stratify

• Do sensitivity analyses

• Don’t expect to find evidence of bias within a meta-analysis

• Summary of findings tables and GRADE offer a convenient way to provide risk of bias assessments alongside results