Critically Appraising Critical Appraisal for Systematic Reviews:
Cochrane Collaboration

Lindsey Anderson
Cochrane
‘Risk of bias’ and ‘quality’

Chapter 8: Assessing risk of bias in included studies

The Cochrane Handbook distinguishes between assessment of methodological quality and RoB

- Assessing RoB targets the extent to which results of included studies should be believed
- A study may be performed to the highest possible standards yet still have an important risk of bias
- Some markers of quality are unlikely to have direct implications for risk of bias
- Assessing RoB overcomes ambiguity between the quality of reporting and the quality of the underlying research
Cochrane

• Cochrane recommend **against** the use of scales yielding a summary score

• Cochrane RoB Tool:
  • A **domain-based evaluation** which assesses RoB in each study with a judgement and a support for the judgement:
  • selection bias, performance bias, detection bias, attrition bias, reporting bias, other bias
Assessments of RoB and synthesis of results (1)

Summary assessments of the RoB for an outcome within each trial should inform the meta-analysis.

Preferred methods:
- restrict the primary meta-analysis to studies at low RoB
- present meta-analyses stratified according to RoB

Higgins et al *BMJ* 2011;343:d5928
Assessments of RoB and synthesis of results (2)

- Meta-regression can be used to compare results from studies at high and low risk of bias, but:
  - comparisons lack power
  - lack of a significant difference should not be interpreted as implying the absence of bias
- Present a meta-analysis of all studies with a summary of the RoB across studies:
  - RoB might be downplayed in the discussion and conclusions, so that decisions continue to be based on flawed evidence
  - This risk could be reduced by assessing the quality of evidence for each important outcome e.g. using the GRADE system

Higgins et al *BMJ* 2011;343:d5928
Quality of the evidence (GRADE)

Cochrane Handbook Chapter 12.2  Assessing the quality of a body of evidence

‘Quality of evidence’ is used in ‘Summary of findings’ tables

- Describes the extent to which you can be confident that an estimate of effect is near the true value for an outcome, across studies
- The RoB of each study is one of several factors that must be considered when judging the quality of a body of evidence
- Highest quality rating is for a body of evidence based on RCTs with low RoB

- Quality assessment can be downgraded on:
  - RoB, imprecision, inconsistency, indirectness, or reporting bias
- Authors judge quality of evidence for each outcome as:
  - ‘High’, ‘Moderate’, ‘Low’, or ‘Very Low’.
- Judgements are made transparent using footnotes or notes
AMSTAR
A Measurement Tool to Assess Systematic Reviews

1. Was an 'a priori' design provided?
2. Was there duplicate study selection and data extraction?
3. Was a comprehensive literature search performed?
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
5. Was a list of studies (included and excluded) provided?
6. Were the characteristics of the included studies provided?
7. Was the scientific quality of the included studies assessed and documented?
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
9. Were the methods used to combine the findings of studies appropriate?
10. Was the likelihood of publication bias assessed?
11. Was the conflict of interest included?

Shea et al. *BMC Medical Research Methodology* 2007, 7:10
Critically Appraising Critical Appraisal for Systematic Reviews: diagnostic test accuracy studies

Harriet Hunt 15/09/2015
What *is* a diagnostic test accuracy study?

- There are a number of ways to assess whether medical tests ‘work’, and one of the most common ways is to measure accuracy
  
  = a test’s ability to discriminate between people *with* the target condition and people *without* the target condition

- This is done by comparing a new test (called the *index test*) against an established ‘best that we’ve got’ test called the *reference standard* and assessing the results using a 2x2 table

<table>
<thead>
<tr>
<th>Index Test</th>
<th>Reference Standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>True Positive</td>
</tr>
<tr>
<td>-</td>
<td>False Negative</td>
</tr>
<tr>
<td></td>
<td>False Positive</td>
</tr>
<tr>
<td></td>
<td>True Negative</td>
</tr>
</tbody>
</table>

www.exeter.ac.uk
Good evidence less easy to come by
Heterogeneity is the norm due to [e.g.]:
Differences between studies in defining a positive test result, test cut-offs and thresholds, study design, patient characteristics, the test’s place in the diagnostic pathway, the purpose of the test...
A highly accurate test does not necessarily improve patient outcome
methodologically complex studies
Relatively new methods on the SR scene
Critical appraisal in diagnostic test accuracy

Tools available:

- Checklists (e.g. the Critical Appraisal Skills Programme (CASP) Diagnostic checklist); use straightforward checklist rating (such as ‘yes’, ‘no’, ‘unsure’) rather than scales or levels of evidence (inappropriate weightings)

- Purpose-built appraisal tools focus on study quality and internal validity (are the observed effects true for people in a study?)
In summary: my experiences

- Very few included studies [3] – still suffered from heterogeneity
- Different thresholds used, most based on the patient group within the study
- Reference standard [PTr] imperfect
- Index test concerns (no explanation of why CA was chosen over other measures)
- Absence of accuracy data on TEG

Conclusions:
- Low quality in test accuracy studies limits applicability
- Critical appraisal particularly important/difficult in this area


Figure 6. Risk of bias and applicability concerns graph: review authors’ judgements about each domain presented as percentages across included studies.
HARRIET HUNT
Research Fellow - Evidence Synthesis and
Test Evaluation
UNIVERSITY OF EXETER MEDICAL SCHOOL
South Cloisters, St Luke's Campus, Exeter EX1 2LU
t +44 (0) 1392 726074
e h.a.hunt@exeter.ac.uk
w http://medicine.exeter.ac.uk