

Systematic reviews of diagnostic accuracy studies 2: Assessment of quality

Mariska Leeflang
Academic Medical Centrum,
University of Amsterdam

Susan Mallett
Centre for Statistics in Medicine,
University of Oxford

Nynke Smidt
Academic Medical Centrum,
University of Amsterdam

With special thanks to:
Hans Reitsma, Anne Rutjes, Vasily Vlassov and Penny Whiting.

Diagnostic Systematic Reviews

- Growing interest:
 - increasing number of published reviews
 - rapid evolution of new diagnostic technologies
- Registration of titles within The Cochrane Collaboration will be able soon!

Cochrane Involvement and Support

- There are currently two Support Units:
 - UK Support Unit (UKSU)
 - Continental Europe Support Unit (CESU)
- Official launch at this Colloquium
- See our website at
<http://srda.cochrane.org/en/index.html>
- For urgent issues and support:
cesu@amc.uva.nl
uksu@contacts.bham.ac.uk

Steps in a systematic review

1. Formulating the question
(and defining criteria for inclusion of studies)
2. Searching for studies
3. Selecting studies and collecting data
4. Assessment of methodological quality
5. Analysing and presenting results
6. Interpreting results

Why assess quality?

- Problem 1: Bias in primary studies can lead to misleading summary estimates of accuracy
- Problem 2: Results of primary studies may vary
- Quality assessment to guide the interpretation of results in terms of potential for bias and sources of heterogeneity

Quality in 114 diagnostic reviews

Method	no (%)
No form of quality assessment	56 (49%)
Table or narrative description	22 (19%)
Inclusion criteria	8 (7%)
Sensitivity analysis	13 (11%)
Meta-regression analysis	7 (6%)

Number of tools: 91

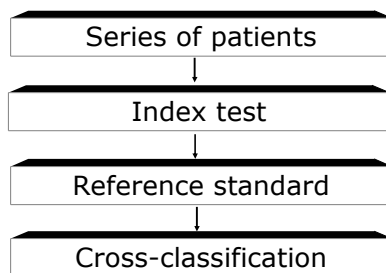
Source: Whiting et al.
Health Technol Assess 2004;8(25)

Cochrane definition of quality

“the methodological quality of a study;
the degree to which the design and
conduct of a study fit to the study
objectives”

Sources of bias and variation

Basic Test Accuracy Study

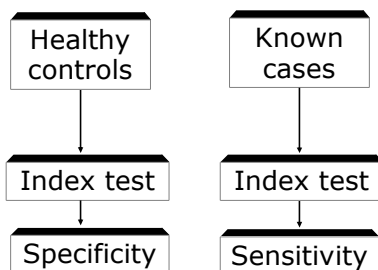


Problems with spectrum

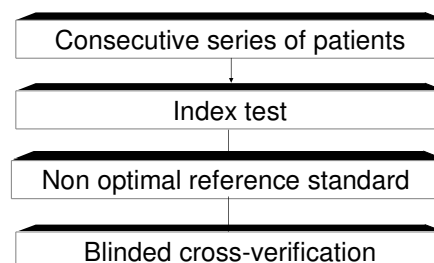
Measures of accuracy vary across
patient groups:

- Patient characteristics e.g. age
- Patient selection/Study design
- Setting

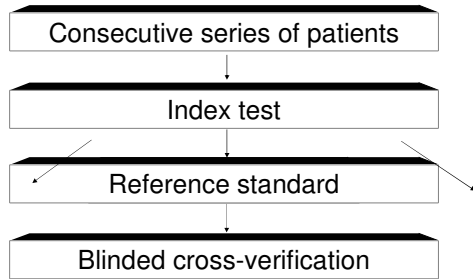
Diagnostic case-control design



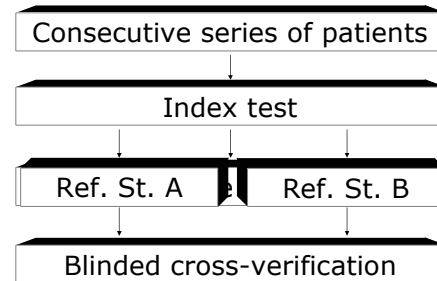
Reference standard bias



Partial verification bias



Differential verification bias



Quality assessment: diagnostic

Many more quality items involved

Whiting P et al. *Ann Intern Med.* 2004 ;140(3):189-202.

Whiting P et al. *BMC Med Res Methodol.* 2003;3:25.

Cochrane Handbook for Reviews of Diagnostic Test Accuracy

Doing the quality assessment

- Quality assessment tools:
 - Large number of different tools
 - Styles: Quality scores/levels of evidence/ component approach
 - The handbook recommends the QUADAS tool for Cochrane reviews

QUADAS

1	Was the spectrum of patients representative of the patients who will receive the test in practice?
2	Is the reference standard likely to correctly classify the target condition?
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?
4	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?
5	Did patients receive the same reference standard regardless of the index test result?
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?
7	Were the index test results interpreted without knowledge of the results of the reference standard?
8	Were the reference standard results interpreted without knowledge of the results of the index test?
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?
10	Were uninterpretable/ intermediate test results reported?
11	Were withdrawals from the study explained?

Additional items

12.	If a cut-off value has been used, was it established before the study was started (pre-specified cut-off value)?
13.	Is the technology of the index test likely to have changed since the study was carried out?
14.	Did the study provide a clear definition of what was considered to be a "positive" result?
15.	Was treatment started after the index test was carried out but before the reference standard was performed?
16.	Was treatment started after the reference standard was carried out but before the index test was performed?
17.	Were data on observer variation reported?
18.	Were data on instrument variation reported?
19.	Were data presented for appropriate patient sub-groups?
20.	Was an appropriate sample size included?
21.	Were objectives pre-specified?

Assessment of items

- All items scored as yes/no/unclear
- Items phrased so that yes indicates absence of bias
- Background document describes how items should be scored

Practical Issues

- Step 1: decide on items to include
- Step 2: develop specific rating guidelines
- Number of assessors: handbook recommends 2
- Background of assessors
- Resolving disagreement
- Piloting the assessment process

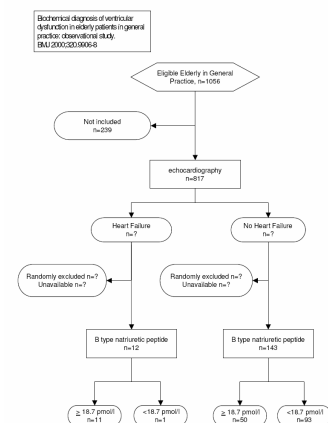
Your quality assessment tool

- 11 recommended QUADAS items
- Should any “additional items” be included?
- Are there any items that are not listed in the table that may be important for your topic area?
- Produce scoring guidelines specific to your review

Now it's your turn!

Example: BNP for heart failure

- **Aim:** To assess the accuracy of BNP for the diagnosis of heart failure
- In small groups:
 1. Produce a flow diagram for the study
 2. Discuss (attention to what has been done, what is missing and possible consequences):
QUADAS item 1 (spectrum)
QUADAS items 2, 4 and 5 (verification)
 3. Discuss the conclusion of the authors



Incorporating study quality

- Present the results of the quality assessment:
 - In a table

	Representative spectrum	Acceptable reference standard	Acceptable delay between tests	Partial verification avoided	Differential verification avoided	Incorporation avoided	Reference standard results blinded	Index test results blinded	Uninterpretable results	Withdrawals	Reported no support from Platelia?
Adam 2004	●	●	?	●	●	●	●	●	●	●	?
Allan 2005	●	?	?	●	●	●	?	?	?	?	?
Becker 2003	●	?	?	●	●	●	●	●	●	●	?
Bretagne 1998	?	?	?	?	?	?	?	?	?	?	?
Buchheit 2004	?	?	?	●	●	●	?	?	?	?	?
Challier 2004	●	?	?	●	●	●	?	?	?	?	?
Doermann 2002	?	?	?	●	●	●	?	?	?	?	?
Fortun 2001	●	?	?	?	?	?	?	?	?	?	?
Herdreht 2002	●	?	?	●	●	●	?	?	?	?	?
Husain 2004	●	●	●	●	●	●	●	●	●	●	?
Machet 1998	?	?	?	●	●	●	?	?	?	?	?
Maertens 2002	?	●	●	●	●	●	?	?	?	?	?
Sulhan 2001	?	?	?	?	?	?	?	?	?	?	?
White 2005	?	?	?	●	●	●	?	?	?	?	?
Williamson 2000	●	?	?	●	●	●	?	?	?	?	?
Yoo 2005	●	?	?	?	?	?	?	?	?	?	?

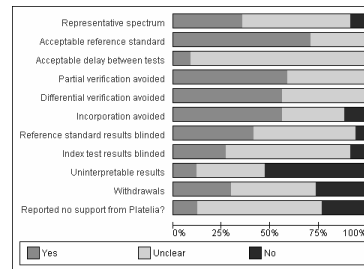
Methodological quality summary.

Review authors' judgments about each methodological quality item for each included study.

Incorporating study quality

- Present the results of the quality assessment:
 - In a table
 - Graphically

Methodological Quality Graph



Review authors' judgments about each methodological quality item presented as percentages across all included studies.

Incorporating study quality

- Present the results of the quality assessment:
 - In a table
 - Graphically
- Investigate individual quality items as potential sources of heterogeneity
- Basis for recommendations for future research

Investigation of heterogeneity

- Stratified analysis according to presence/absence of specific quality criteria
- Several features simultaneously using meta-regression analysis
- Define methodological criteria a priori

Problems with quality assessment

- Not as straightforward as it might sound!
- Hampered by poor reporting
- Quality assessment is subjective
- Statistical incorporation of quality problematic with limited studies

Conclusions

- Quality assessment is essential
- The QUADAS tool should be used as a starting point
- The quality items and scoring guidelines should be tailored to your review question
- The results of the quality assessment should be presented
- No quality scores and cut-offs for 'good' quality
- Study quality should be incorporated into all reviews

Questions?