Centre for Reviews and Dissemination

The development of QUADAS:

a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews

Penny Whiting MSc[1], Anne W.S. Rutjes MSc[2], Johannes B. Reitsma MD PhD[2], Patrick M.M. Bossuyt PhD[2], Jos Kleijnen MD PhD[1]

[1]Centre for Reviews and Dissemination, The University of York

[2]Department of Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam

1. Review of methodological literature

- 55 studies that investigated the effects of bias and variation on measures of test performance were included
- Sources of bias supported by the most empirical evidence:
  - variation by clinical and demographic subgroups
  - disease prevalence/severity
  - partial verification bias
  - clinical review bias
  - observer/instrument variation.

2. Review of how quality assessment has been used in systematic reviews

- 58 reviews of diagnostic accuracy studies, that conducted some form of quality assessment, identified from DARE were included
- Methods of incorporating quality:
  - narrative synthesis (43)
  - presented in a table (37)
  - recommendations for future research (11)
  - inclusion criteria (8)
  - inclusion in primary analysis (7)
  - to conduct sensitivity analyses (13)
  - to quantify quality related variation in regression analyses (7)

3. Review of tools used to assess the quality of diagnostic tests

- 92 quality assessment tools were included
- None had been validated and only two provided details of development
- Review bias was the only item included in more than 75% of tools. Four items were included in more than half:
  - spectrum composition
  - population recruitment
  - the use of an appropriate reference standard
  - verification bias

Item generation: initial list of items

- An initial list of 28 items was produced

Preliminary conceptual decisions

- The tool was required to:
  - be used in systematic reviews of diagnostic accuracy
  - assess the methodological quality of a diagnostic study
  - allow consistent and reliable assessment of quality
  - be short and simple to complete
- 'Quality' included internal and external validity
- Had to distinguish high from low quality studies
- Component analysis was selected as the best approach
- No quality score

Assessment of face validity: Delphi procedure

- Nine experts participated
- A summary of the evidence from the three reviews was provided
- 28 items rated for inclusion
- Asked to suggest additional items or rephrasing of items
- Four rounds, after which agreement was reached on final items

QUADAS

Item		Yes	No	Unclear
1.	Was the spectrum of patients representative of the patients who will receive the test in practice?			
2.	Were selection criteria clearly described?			
3.	Is the reference standard likely to classify the target condition correctly ?			
4.	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?			
5.	Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?			
6.	Did patients receive the same reference standard regardless of the index test result?			
7.	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?			
8.	Was the execution of the index test described in sufficient detail to permit its replication?			
9.	Was the execution of the reference standard described in sufficient detail to permit its replication?			
10.	Were the index test results interpreted without knowledge of the results of the reference standard?			
11.	Were the reference standard results interpreted without knowledge of the results of the index test?			
12.	Were the clinical data available when test results were interpreted the same as when the test would be used in practice?			
13.	Were uninterpretable or intermediate test results reported?			
14.	Were withdrawals from the study explained?			

Field trials to assess consistency and construct validity

- Evaluation of QUADAS is in progress:
  - piloting in published studies, focussing on consistency and reliability
  - piloting in diagnostic reviews
  - regression analysis to investigate associations between quality and diagnostic accuracy

Generation of the refined instrument

- QUADAS will be adapted as necessary

Further development of QUADAS

- Development of design & topic specific criteria
- Possible areas: screening, clinical examination, biochemical tests, imaging evaluations, invasive procedures, genetic markers, diagnostic case control studies and diagnostic cohort studies

Completed stages

Ongoing stages

This work was commissioned and funded by the NHS R&D Health Technology Assessment programme, project number 98/27/99