Methods for assessing the quality of diagnostic accuracy studies

Development and validation of methods for assessing the quality of diagnostic accuracy studies

P Whiting,[1*] AWS Rutjes,[2] J Dinnes,[3] JB Reitsma,[2] PMM Bossuyt[2] and J Kleijnen[1]

[1] Centre for Reviews and Dissemination, University of York, UK
[2] Department of Clinical Epidemiology and Biostatistics, Academic Medical Centre, University of Amsterdam, The Netherlands
[3] Wessex Institute for Health Research and Development, University of Southampton, UK

* Corresponding author

Executive summary

Health Technology Assessment 2004; Vol. 8: No. 25

Health Technology Assessment
NHS R&D HTA Programme

Executive summary

Background

The assessment of the quality of studies included in a systematic review is as important for reviews of studies of diagnostic accuracy as it is for any other type of review. There is currently a lack of a validated tool for the assessment of such studies.

Objectives

This project aims to develop a quality assessment tool which will be used in systematic reviews to assess the quality of primary studies of diagnostic accuracy.

Methods

Three systematic reviews were conducted to provide an evidence base for the development of the quality assessment tool. The methodological literature on diagnostic test assessment was reviewed to identify potential sources of bias. Systematic reviews of diagnostic tests that used any form of quality assessment were examined to identify how quality was incorporated. Lastly, a review of existing quality assessment tools was conducted to ascertain what methods exist for assessing the quality of diagnostic studies, and on what evidence they are based. Literature searches were used to identify studies for each of the reviews. Systematic inclusion criteria were applied; studies were selected for relevance and inclusion by one reviewer and checked by a second. Data for each of the reviews were extracted into an Access database by one reviewer and checked by a second. All discrepancies were resolved by discussion or through consultation with a third reviewer when agreement could not be reached. A narrative synthesis is presented for each of the reviews.

A Delphi procedure was used to develop the quality assessment tool. The information provided by the reviews was incorporated into this. A panel of nine experts in the area of diagnostic accuracy studies took part in the Delphi procedure. In the first round members were asked to indicate which of the items on the initial list of items (provided by the results of the reviews) should be included in the tool. Items for which there were high levels of agreement were selected for inclusion/exclusion in the tool; items for which there was disagreement were rated again as part of the next round. Panel members were also asked to make comments and to suggest rephrasings of the items or additional items if appropriate. During subsequent rounds the results of previous rounds were fed back to panel members and they were asked to rerate the items based on the results of the previous rounds. The procedure was continued until agreement was reached on which items were to be included in the quality assessment tools. Panel members were also asked to provide feedback on various other items such as the proposed scoring method, whether they endorsed the procedure, whether they had used the evidence provided to them, and whether they would like to see the development of additional topic and design specific items.

The Delphi procedure produced the quality assessment tool, named QUADAS. A background document was produced which gives details on what is meant by each item included in the tool and how each of the items should be scored.

Work to validate the tool will continue beyond the scope of this project. The validation process will include the piloting of the tool on a small sample of published studies, assessment of the consistency and reliability of the tool, piloting the tool in a number of diagnostic reviews, and using a regression analysis to investigate associations between study characteristics and estimates of diagnostic accuracy in primary studies, as combined in existing systematic reviews.

Results

The reviews produced a list of 28 possible items for inclusion in the quality assessment tool. The first review found that the sources of bias supported by the most empirical evidence were variation by clinical and demographic subgroups, disease prevalence/severity, partial verification bias, clinical review bias and observer/instrument variation. There was also some evidence of bias for

the effects of distorted selection of participants, absent or inappropriate reference standard, differential verification bias and review bias. The evidence for the effects of other sources of bias was insufficient to draw conclusions regarding the effects, if any, of these biases. The third review found that only one item, the avoidance of review bias, was included in more than 75% of tools. A further four items were each included in 50–75% of tools: spectrum composition, population recruitment, absent or inappropriate reference standard and verification bias. Other items were included in less than 50% of tools.

The second review found that the quality assessment tool needs to have the potential to be discussed narratively, reported in a tabular summary, used as recommendations for future research, used to conduct sensitivity or regression analyses and used as criteria for inclusion in the review or a primary analysis. The resulting implication for the development of the tool is that some distinction needs to be made between high- and low-quality studies. It was decided that component analysis is the best approach to incorporate quality into systematic reviews of diagnostic studies. The quality tool was developed taking this into consideration.

The Delphi procedure consisted of four rounds, after which agreement was reached on the items to be included in QUADAS. The final tool included 14 items:

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 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 replication of the test?
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 same clinical data available when test results were interpreted as would be available when the test is used in practice?
13. Were uninterpretable/intermediate test results reported?
14. Were withdrawals from the study explained?

Conclusions

This project produced an evidence-based quality assessment tool to be used in systematic reviews of diagnostic accuracy studies. Through the various stages of the project the current lack of such a tool and the need for a systematically developed validated tool were demonstrated. Further work to validate the tool continues beyond the scope of this project. The further development of the tool by the addition of design- and topic-specific criteria is proposed.

Publication

Whiting P, Rutjes AWS, Dinnes J, Reitsma JB, Bossuyt PMM, Kleijnen J. Development and validation of methods for assessing the quality of diagnostic accuracy studies. *Health Technol Assess* 2004;**8**(25).

NHS R&D HTA Programme

The research findings from the NHS R&D Health Technology Assessment (HTA) Programme directly influence key decision-making bodies such as the National Institute for Clinical Excellence (NICE) and the National Screening Committee (NSC) who rely on HTA outputs to help raise standards of care. HTA findings also help to improve the quality of the service in the NHS indirectly in that they form a key component of the 'National Knowledge Service' that is being developed to improve the evidence of clinical practice throughout the NHS.

The HTA Programme was set up in 1993. Its role is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The HTA programme commissions research only on topics where it has identified key gaps in the evidence needed by the NHS. Suggestions for topics are actively sought from people working in the NHS, the public, consumer groups and professional bodies such as Royal Colleges and NHS Trusts.

Research suggestions are carefully considered by panels of independent experts (including consumers) whose advice results in a ranked list of recommended research priorities. The HTA Programme then commissions the research team best suited to undertake the work, in the manner most appropriate to find the relevant answers. Some projects may take only months, others need several years to answer the research questions adequately. They may involve synthesising existing evidence or designing a trial to produce new evidence where none currently exists.

Additionally, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme is able to commission bespoke reports, principally for NICE, but also for other policy customers, such as a National Clinical Director. TARs bring together evidence on key aspects of the use of specific technologies and usually have to be completed within a limited time period.

.Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned by the HTA Programme as project number 98/27/99. As funder, by devising a commissioning brief, the HTA Programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health.

HTA Programme Director: Professor Tom Walley
Series Editors: Professor John Gabbay, Dr Chris Hyde, Dr Ruairidh Milne, Dr Rob Riemsma and Dr Ken Stein
Managing Editors: Sally Bailey and Caroline Ciupek

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2004

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to NCCHTA, Mailpoint 728, Boldrewood, University of Southampton, Southampton, SO16 7PX, UK.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.
Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.