**Review**

*Executive summary*

# Assessing the quality of reports of randomised trials: implications for the conduct of meta-analyses

D Moher[1,2]    M Moher[6]
DJ Cook[3]      A Jones[1]
AR Jadad[4]     B Pham[1]
P Tugwell[5]    TP Klassen[1,2]

[1]  Thomas C Chalmers Centre for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, Canada
[2]  Department of Pediatrics, University of Ottawa, Canada
[3]  Department of Clinical Epidemiology and Biostatistics, McMaster University, Canada
[4]  Health Information Research Unit, Department of Clinical Epidemiology and Biostatistics, McMaster University, Canada
[5]  Department of Medicine, University of Ottawa, Canada
[6]  Division of Public Health and Primary Health Care, Institute of Health Sciences, University of Oxford, UK

**Health Technology Assessment
NHS R&D HTA Programme**

HTA
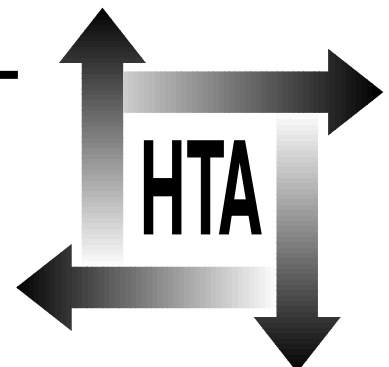
# Executive summary

## Objectives

- To examine the issue of quality assessment of randomised controlled trials (RCTs) included in meta-analyses.
- To provide empirically based recommendations on how to conduct meta-analyses with respect to quality assessment.

Five projects were carried out to achieve these objectives.

1. A database of meta-analyses was developed that provided the majority of data for the remaining projects.
2. Journal editors, methodologists and systematic reviewers associated with randomly selected articles in the database were surveyed about their views on the assessment and reporting of quality of the primary trials included in meta-analyses.
3. The frequency of quality assessment and the methods used were investigated using a sample of meta-analyses (n = 240) from the main database.
4. The effect that the quality of RCTs included in a meta-analysis has on estimates of intervention effectiveness was analysed using a sample of meta-analyses (n = 11 covering 127 RCTs) from the database.
5. Guidelines were developed on the basis of the evidence obtained in the other projects.

## Data sources

A comprehensive list of studies was provided by an electronic search of databases including MEDLINE, EMBASE and the Cochrane Database of Systematic Reviews (CDSR).

## Study selection

Meta-analyses were selected. The inclusion criterion was that the study combined (pooled) the overall results of RCTs included in the meta-analysis.

## Data extraction

Data extraction forms were used to extract the necessary data from the articles. Data extraction was completed in duplicate to reduce the chances of error. Inter-rater reliability was calculated before data extraction began.

## Data synthesis

Quantitative analysis was difficult because of the nature of the research questions and was conducted only for the study examining the effect of RCT quality on estimates of intervention effectiveness. The data for the searching study and the survey, and the descriptive data for the quality assessment study, are discussed mostly in a qualitative manner.

## Results

The overlap of articles and journals between MEDLINE and EMBASE was 80% and 87%, respectively. The database of 491 articles that was used comprised 455 meta-analyses identified by the MEDLINE search and 36 meta-analyses in the CDSR.

Response rates from the survey were 78%, 74% and 59% for reviewers (n = 121), methodologists (n = 55) and editors (n = 63), respectively. Over 90% of respondents stated that assessment and reporting of quality of RCTs included in meta-analyses was very or somewhat important. The use of RCT design features such as inclusion criteria, and using quality assessments to conduct sensitivity analyses were the most frequently endorsed methods of incorporating the quality assessments into meta-analyses. Most respondents believed that guidelines on the assessment and reporting of the quality of randomised trials would increase the rigour of reporting of published meta-analyses and make interpretation easier.

Of a sample of 240 meta-analyses, trial quality was assessed in 48% and in half of these data on the reproducibility of the assessments were provided. Of the meta-analyses that assessed quality, only 25% incorporated trial quality into the analyses.

Masked and unmasked quality assessments were carried out on 127 RCTs included in 11 meta-analyses in the database. The assessments were made using a validated scale (1–5, higher scores indicate superior reporting) and individual components known to affect estimates of intervention effectiveness. Masked quality assessment provided significantly higher scores (mean = 2.74;

standard deviation (SD) = 1.10) than unmasked assessments (mean = 2.55; SD = 1.20). Low-quality trials were associated with an increase of 34% in estimate of benefit (ratio of odds ratios (ROR) = 0.66; 95% confidence interval (CI): 0.52, 0.83) compared with high-quality trials. Trials using inadequate allocation concealment, compared with those using adequate methods, were also associated with an increased estimate of benefit of 37% (ROR = 0.63; 95% CI: 0.45, 0.88). The average treatment benefit across all trials was 39% (OR = 0.61; 95% CI: 0.57, 0.65). Including only trials with low quality scores increased this effect to 52% (OR = 0.48; 95% CI: 0.43, 0.54), whereas including only trials with high quality scores reduced the effect to 29% (OR = 0.71; 95% CI: 0.65, 0.77). Using all the trial scores as quality weights reduced the effect to 35% (OR = 0.65; 95% CI: 0.59, 0.71) and resulted in the least statistical heterogeneity.

## Conclusions

Indexing inconsistencies within and across databases pose challenges in searching for systematic reviews or meta-analyses. Our results suggest that it is necessary to search multiple databases to identify all relevant information. Journal indexers, authors and editors should collaborate to develop and implement criteria to help users of systematic reviews and meta-analyses identify relevant publications.

The systematic reviewers, methodologists and journal editors surveyed believed that assessment of trial quality was important. This contrasts with the infrequent reporting of trial quality in published meta-analyses. Future studies should address the issue of quality assessment. Consistent reporting of the design features of RCTs may help to enhance the rigour and clinical interpretability of meta-analyses.

Among a sample of meta-analyses from the database, individual components and scales were the methods most commonly used to assess trial quality. However when quality assessments were made, in most cases they were not incorporated into the analysis. This is important because the incorporation of quality assessments can alter the estimate of the benefit of intervention, regardless of which method of assessment is used.

The results from these studies also suggest that certain characteristics of the design and execution of RCTs impact on the probability of bias, and further research is needed on this. Investigations are also needed to clarify the value of masking studies before quality assessment and to determine the advantages of the various approaches to incorporate quality assessments into the analyses. Until such empirical evidence is presented, the guidelines outlined below are a useful tool with which meta-analysts, editors, peer reviewers and readers can deal with issues pertaining to quality assessment of randomised trials included in a meta-analysis.

## Guidelines

- The quality of all randomised trials included in a meta-analysis should be assessed.
- Masked quality assessment should be considered, and meta-analysts should report masking methods used or their reasons for rejecting masking.
- Primarily evidence-based components (e.g. allocation concealment, double-blinding, type of randomised trial) should be used to assess quality. Topic-specific items should be part of the quality-assessment process.
- Scales used for assessment should have been appropriately developed and evaluated. A component approach has the advantage that it can be topic-specific. However, there is no compelling evidence to recommend a component approach over a scale approach or vice versa.
- Meta-analyses should incorporate an estimate of quality assessment into the quantitative analysis as a 'first-line' sensitivity analysis.

## Publication

# NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme 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 work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Methodology Panel and funded as project number 93/52/04.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

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.

Series Editors:        Andrew Stevens, Ruairidh Milne and Ken Stein
Editorial Assistant:    Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.