

How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study

M Egger^{1,2*}

P Juni^{2,3}

C Bartlett²

F Holenstein¹

J Sterne²

¹ Department of Social and Preventive Medicine, University of Berne, Switzerland

² Medical Research Council Health Services Research Collaboration, Department of Social Medicine, University of Bristol, UK

³ Department of Rheumatology and Clinical Immunology, University of Berne, Switzerland

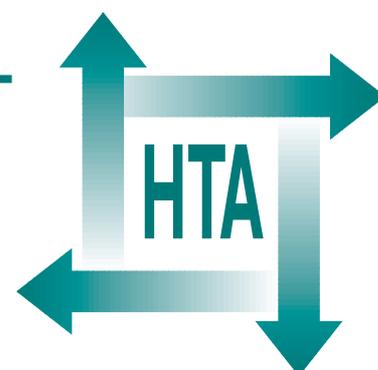
* Corresponding author



Executive summary

Health Technology Assessment 2003; Vol. 7: No. 1

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





Executive summary

Background

The inclusion of an unbiased sample of relevant studies is central to the validity of systematic reviews and meta-analyses. Time-consuming and costly literature searches, which cover the grey literature and all relevant languages and databases, are normally recommended to prevent reporting biases. However, the size and direction of these effects is unclear at present. There may be trade-offs between timeliness, cost and the quality of systematic reviews.

Objectives

- To examine the characteristics of clinical trials that are difficult to locate (unpublished trials, trials published in languages other than English, trials published in journals not indexed in the MEDLINE database) and of trials of lower quality (inadequate/unclear concealment of treatment allocation, not double-blind).
- To compare within meta-analyses the treatment effects reported in trials that are difficult to locate with trials that are more accessible, and of trials of lower with trials of higher quality.
- To assess the impact of excluding trials that are difficult to locate and of trials of lower quality on pooled effect estimates, p -values and the shape of funnel plots.

Methods

Data sources

The following sources were searched for relevant meta-analyses:

- eight medical journals that regularly publish systematic reviews (handsearch)
- systematic reviews published in the Cochrane Database of Systematic Reviews
- systematic reviews included in the Database of Abstracts of Reviews of Effectiveness
- *Health Technology Assessment* (handsearch).

Study selection

Meta-analyses of therapeutic or preventive interventions that were based on comprehensive literature searches and which combined the binary

outcomes of at least five controlled clinical trials were included. Comprehensive literature searches were defined as follows:

- the search was not restricted to the English language literature
- the Cochrane Controlled Trials Register or at least two other electronic databases (such as MEDLINE or EMBASE) had been searched
- at least one indicator of searches for unpublished trials was present (e.g. searches of conference proceedings or contacts with licensing bodies).

Data extraction

Trial reports were classified as published journal articles if they had been published as full or short reports, editorials or letters in a regular or supplementary issue of a journal. Language was assessed using the SERLINE journals database, and published trials were classified according to whether or not they had been published in a MEDLINE-indexed journal. Quality assessment was restricted to trials included in Cochrane reviews.

Data synthesis

Meta-analyses that were able to contribute to the analysis in question were included. For example, only meta-analyses that contained both published and unpublished trials were included in the analyses addressing the impact of publication bias. Within each meta-analysis pooled effect estimates were calculated separately for the trials that are difficult to locate and the remaining trials, applying the same statistical model used by the original authors. For each meta-analysis a ratio of the pooled estimates was derived. A weighted average for all these ratios was calculated using random-effects meta-analysis. The percentage change in the pooled effect estimate which occurred when trials that are difficult to locate were excluded, was also calculated and changes in p -values and the impact on the shape of the funnel plot (using a regression method to measure funnel plot asymmetry) were examined.

Results

- A total of 159 systematic reviews met the inclusion criteria but not all included trials that are

difficult to locate. Comparisons of treatment effects were based on the following:

- unpublished versus published (60 meta-analyses)
- other languages versus English (50 meta-analyses)
- non-indexed versus MEDLINE-indexed (66 meta-analyses).

Analyses of trial quality were based on:

- inadequately concealed/unclear versus adequately concealed (39 meta-analyses)
- not double-blind versus double-blind (45 meta-analyses).

- The importance of trials that are difficult to locate appears to vary across medical specialities. For example, unpublished trials are particularly prevalent in oncology whereas trials published in languages other than English and trials published in sources not indexed in MEDLINE are important in psychiatry, rheumatology and orthopaedics. A large proportion of trials of complementary medicine are difficult to locate.
- Unpublished trials show less beneficial effects than published trials whereas non-English language trials and non-indexed trials tend to show larger treatment effects.
- Trials that are difficult to locate tend to be smaller and of lower methodological quality than trials that are easily accessible and published in English.
- Trials with inadequate or unclear concealment of allocation show more beneficial effects than adequately concealed trials. Similarly, open trials tend to be more beneficial than double-blind trials.
- In the majority of meta-analyses exclusion of trials with inadequate or unclear concealment and trials without double-blinding led to a change towards less beneficial treatment effects, which was often substantial.
- Including unpublished trials reduces funnel plot asymmetry whereas the inclusion of trials published in languages other than English and of non-indexed trials increases the degree of asymmetry in the funnel plot. The impact of trials of lower methodological quality on the funnel plot is substantial for trials with inadequate or unclear concealment of allocation.

Conclusions

Systematic reviews that are based on a search of English language literature that is accessible in the

major bibliographic databases will often produce results that are close to those obtained from reviews based on more comprehensive searches that are free of language restrictions. We recommend that when planning a review, investigators should consider the type of literature search and the degree of comprehensiveness that are appropriate for the review in question, taking into account budgetary and time constraints.

The finding that trials which are difficult to locate are often of lower quality raises the worrying possibility that rather than preventing bias through extensive literature searches, bias could be introduced by including trials of low methodological quality. We believe that in situations where resources are limited, thorough quality assessments should take precedence over extensive literature searches and translations of articles.

Our results confirm that the funnel plot and the regression method to assess funnel plot asymmetry are useful to detect 'small-study effects', the tendency for smaller studies in a meta-analysis to show larger treatment effects.

Recommendations for future research

- The importance of trials that are difficult to locate appears to vary not only between conventional and complementary medicine but also within conventional medicine. Further research is required to clarify this issue.
- Future studies should prospectively compare the results from rapid reviews that are restricted to the English language with meta-analyses based on extensive searches without language restrictions.
- The inclusion or exclusion of trials of low methodological quality has a substantial impact on results and conclusions from systematic reviews and meta-analyses. Further methodological research into markers of trial quality in different areas of medicine is required.

Publication

Egger M, Jüni P, Bartlett C, Hoenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess* 2003;**7**(1).

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 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.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 97/18/05.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme 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 any recommendations made by the authors.

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.

Methodology Programme Director: Professor Richard Lilford
HTA Programme Director: Professor Kent Woods
Series Editors: Professor Andrew Stevens, Dr Ken Stein, Professor John Gabbay,
Dr Ruairidh Milne and Dr Chris Hyde
Managing Editors: Sally Bailey and Sarah Llewellyn Lloyd

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.