Finding evidence to inform the conduct and reporting of systematic reviews of randomized trials

Moher, D.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement

David Moher, Deborah J. Cook, Susan Eastwood, Ingram Olkin, Drummond Rennie, Donna F. Stroup, for the QUOROM Group

Lancet 1999; 354:1896-1900
Chapter 10

Summary

**Background:** The QUality Of Reporting Of Meta-analyses (QUOROM) conference was convened to address standards for improving the quality of reporting of meta-analyses of clinical randomized controlled trials (RCTs).

**Methods:** The QUOROM group comprised 30 clinical epidemiologists, clinicians, statisticians, editors and researchers. In conference, the group was asked to identify items they thought should be included in a checklist of standards. Whenever possible, checklist items were guided by research evidence suggesting that failure to adhere to the item proposed could lead to biased results. A modified Delphi technique was used in assessing candidate items.

**Findings:** The conference resulted in the QUOROM statement, a checklist and a flow diagram. The checklist describes an optimal way to present the Abstract, Introduction, Methods, Results and Discussion sections of a report of a meta-analysis. It is organized into 18 headings and subheadings in the Methods section regarding searches, selection, validity assessment, data abstraction, study characteristics, and quantitative data synthesis, and in the Results with study characteristics and quantitative data synthesis; research documentation was identified for nine of the 18 items. The flow diagram provides information about both the number of RCTs identified, included, and excluded and the reasons for excluding trials.

**Interpretation:** We hope this report will generate further thought about ways to improve the quality of reports of meta-analyses of RCTs and that interested readers, reviewers, researchers and editors to use the QUOROM statement and generate ideas for improvement.
Introduction

Health-care providers and other decision makers now have, among their information resources, a form of clinical report called the meta-analysis—a review in which bias has been reduced by the systematic identification, appraisal, synthesis, and, if relevant, statistical aggregation of all relevant studies on a specific topic according to a predetermined and explicit methodology. The number of published meta-analyses has increased substantially in the past decade. Not only can these integrative articles be helpful for clinical decisions, but they may also serve as the policy foundation for evidence-based practice guidelines, economic evaluations, and future research agendas. The value of meta-analysis is evident in the work of the international Cochrane Collaboration, whose primary purpose is to generate and disseminate high-quality systematic reviews of health care interventions.

Like any research enterprise, particularly one that is observational, the meta-analysis of evidence can be flawed. Accordingly, the process by which meta-analyses are conducted has recently undergone scrutiny. A 1987 survey of 86 English-language meta-analyses assessed each publication on 23 items from six content areas considered important in the conduct and reporting of a meta-analysis of randomized trials: study design, combinability, control of bias, statistical analysis, sensitivity analysis, and problems of applicability. The survey results indicated that only 24 (28%) of the 86 meta-analyses reported addressing all six content areas. The updated survey, which included more recently published meta-analyses, showed little improvement in the rigour with which they were reported.

Several publications have described the science of reviewing research, differences among narrative reviews, systematic reviews, and meta-analyses, and how to conduct, critically appraise, and apply meta-analyses in practice. The burgeoning number of meta-analyses published has highlighted such issues as discordant meta-analyses on the same topic and discordant meta-analyses and randomized-trial results on the same question.

An important consideration in interpreting and using meta-analyses is to ascertain that the investigators who performed the meta-analysis not only report explicitly the methods they used to analyze the articles they reviewed, but also report the methods used in the research articles they analyzed. The meta-analytic review methods used may not be provided in an initial manuscript submission: even when they are, other factors such as page limitations, peer review and editorial decisions, may alter the content and format of the manuscript before publication.

Several authors have suggested guidelines for reporting meta-analyses. However, a consensus across disciplines about how meta-analyses should be
reported has not been developed. Following a recent initiative to improve the quality of reporting of randomized controlled trials,\textsuperscript{20-22} we organized the Quality Of Reporting Of Meta-analyses (QUOROM) conference to address these issues as they relate to meta-analyses of randomized trials. This report summarizes the proceedings of that conference. The issues discussed might also be useful for reporting of systematic reviews (i.e., meta-analysis, as defined above, without statistical aggregation), particularly of randomized trials.

**Methods**

The QUOROM steering committee began with a comprehensive review of the published literature on the conduct and reporting of meta-analyses. The databases searched included MEDLINE and The Cochrane Library,\textsuperscript{23} which consists of the Cochrane Database of Systematic Reviews (CDSR), the Cochrane Controlled Trials Register (CCTR), the York Database of Abstracts of Reviews of Effectiveness (DARE), and the Cochrane Review Methodology Database (CRMD). We examined reference lists of the retrieved articles and individual personal files. Articles of potential relevance were retrieved and critically appraised by the QUOROM steering committee. The committee generated a draft agenda for the conference, which included six domains requiring discussion and debate. The content areas were slightly modified during preliminary discussions at the conference and are reported as: 1) the search for the evidence; 2) decision making regarding which evidence to include; 3) description of the characteristics of primary studies; 4) quantitative data synthesis; 5) reliability and issues related to internal validity (or quality); and 6) clinical implications related to external validity. (or generalizability)

In planning the QUOROM conference, the steering committee identified clinical epidemiologists, clinicians, statisticians, and researchers who conduct meta-analysis as well as editors from the United Kingdom and North America who were interested in meta-analysis. These 30 individuals were invited to a two-day conference in Chicago on October 2-3, 1996. Participants were surveyed before the meeting to elicit their views regarding current reporting standards of meta-analyses and whether reporting standards needed improvement. In addition, they were sent relevant citations for review and were asked to indicate in which of the six groups they wished to participate.

The conference format comprised small group and plenary sessions. Each small group had a facilitator who was a member of the steering committee and responsible for ensuring the discussions of as many of the issues relevant to their specific remit as possible. Each small group also had a recorder, who was responsible for documenting the main points and the consensus regarding each issue discussed during that session; the recorder presented the group's consensus
during the plenary sessions. During the plenary sessions, an elected scribe from each small group was responsible for recording the principle points relevant to that group's charge that arose during the plenary discussion.

The participants in each small group were asked to identify items that they thought should be included in a checklist of standards that would be useful for investigators, editors, and peer reviewers. We asked that, whenever possible, items included in the checklist be guided by research evidence suggesting that a failure to adhere to the particular checklist item proposed could lead to biased results. For example, a substantial lack of sensitivity and specificity of MEDLINE searches is evident. Therefore, the checklist suggests that investigators explicitly describe all search strategies used to locate articles for inclusion in a meta-analysis. In considering whether candidate items were essential, each subgroup used a modified Delphi technique that was replicated in the plenary sessions.

Results

The conference resulted in the QUOROM statement: a checklist (Table 1) and a flow diagram (Figure 1). The checklist of standards for reporting meta-analyses describes an optimal way to present the Abstract, Introduction, Methods, Results and Discussion sections of a report of a meta-analysis. The checklist is organized into 18 headings and subheadings to encourage authors to provide readers with information regarding searches, selection, validity assessment, data abstraction, study characteristics, quantitative data synthesis, and trial flow. For trial flow, authors are asked to provide a flow diagram (See Figure 1) providing information about both the number of randomized trials identified, included, and excluded along with the reasons for excluding them.
Table 1. Quality of reporting of meta-analyses—QUORUM for clinical randomized controlled trials (RCTs)

<table>
<thead>
<tr>
<th>Heading</th>
<th>Subheading</th>
<th>Descriptor</th>
<th>Reported? [Y/N]</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td></td>
<td>Identify the report as a meta-analysis [or systematic review¹] of randomized trials (RCTs)²⁶</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td>Use a structured format²⁷ Describe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Objectives</td>
<td></td>
<td>The clinical question explicitly.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data sources</td>
<td></td>
<td>The databases [i.e., list] and other information sources.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review methods</td>
<td></td>
<td>The selection criteria [i.e., population, intervention, outcome, and study design]; methods for validity assessment, data abstraction, and study characteristics, and quantitative data synthesis] in sufficient detail to permit replication.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td>Characteristics of the randomized trials included and excluded; qualitative and quantitative findings [i.e., point estimates and confidence intervals]; and subgroup analyses.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conclusion</td>
<td></td>
<td>The main results.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text Introduction</td>
<td></td>
<td>Describe The explicit clinical problem, biologic rationale for the intervention, and rationale for review.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 1 (continued). Quality of reporting of meta-analyses—QUORUM for clinical randomized controlled trials (RCTs)

<table>
<thead>
<tr>
<th>Heading</th>
<th>Subheading</th>
<th>Descriptor</th>
<th>Reported?</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Text</td>
<td>Describe</td>
<td>The information sources, in detail(^\text{28}) [e.g., databases, registers, personal files, expert informants, agencies, hand-searching], and any restrictions [years considered, publication status,(^\text{29}) language of publication].(^\text{30,31})</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>Searching</td>
<td>The inclusion and exclusion criteria [defining population, intervention principal outcomes, and study design].(^\text{32})</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Selection</td>
<td>The criteria and process used [e.g., masked conditions, quality assessment and their findings].(^\text{33,34,35,36})</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validity</td>
<td>Data abstraction</td>
<td>The process or processes used [e.g., completed independently, in duplicate].(^\text{35,36})</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>assessment</td>
<td>Study characteristics</td>
<td>The type of study design, participants' characteristics, details of intervention, outcome definitions, etc.,(^\text{37}) and how clinical heterogeneity was assessed.</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>Quantitative</td>
<td>The principal measures of effect [e.g., relative risk], method of combining results [statistical testing and confidence intervals], handling of missing data, etc.; how statistical heterogeneity was assessed,(^\text{38}) a rationale for any a priori sensitivity and subgroup analyses; and any assessment of publication bias.(^\text{39})</td>
<td>N</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# Chapter 10

## Table 1 (continued). Quality of reporting of meta-analyses—QUORUM for clinical randomized controlled trials (RCTs)

<table>
<thead>
<tr>
<th>Heading</th>
<th>Subheading</th>
<th>Descriptor</th>
<th>Reported? [Y/N]</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Results</td>
<td>Trial flow</td>
<td>Provide a meta-analysis profile summarizing trial flow [see figure].</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Study characteristics</td>
<td>Present descriptive data for each trial [e.g., age, sample size, intervention, dose, duration, follow-up period].</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quantitative data synthesis</td>
<td>Report agreement on the selection and validity assessment; present simple summary results [for each treatment group in each trial, for each primary outcome]; present data needed to calculate effect sizes and confidence intervals in intention-to-treat analyses [e.g., $2 \times 2$ tables of counts, means and standard deviations, proportions].</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
<td>Summarize the key findings; discuss clinical inferences based on internal and external validity; interpret the results in light of the totality of available evidence; describe potential biases in the review process [e.g., publication bias]; and suggest a future research agenda.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

186
Figure 1. Progress through the stages of a meta-analysis, including selection of potentially relevant randomized controlled trials [RCTs], included and excluded RCTs with a statement of the reasons, RCTs with usable information, and RCTs withdrawn by outcome with a statement of the reasons for the withdrawal.

1. Potentially relevant RCTs identified and screened for retrieval [n = ...]

2. RCTs excluded, with reasons [n = ...]

3. RCTs retrieved for more detailed evaluation [n=...]

4. RCTs excluded, with reasons [n = ...]

5. Potentially appropriate RCTs to be included in the meta-analysis [n=...]

6. RCTs included in meta-analysis [n = ...]

7. RCTs excluded from the meta-analysis, with reasons [n = ...]

8. RCTs with usable information, by outcome [n = ...]
Pre-testing of checklist and flow diagram

After development of the checklist and flow diagram, two members of the steering committee (DM, DJC) conducted pre-testing with epidemiology graduate students studying meta-analysis, general internal medicine residents, participants attending a Canadian Cochrane Center workshop, and faculty members of Departments of Medicine and of Epidemiology and Biostatistics. One group of candidates for a Master's degree in Epidemiology used the checklist and flow diagram to report their meta-analyses as if their work were being submitted for publication. Feedback from these four groups was positive, most users stating that the QUOROM checklist and flow diagram would be likely to improve reporting standards. Modifications of the checklist (e.g., inclusion of a statement about major findings) and changes to the flow diagram (e.g., more detail) were incorporated and are reflected in the Table and Figure for this report.

Discussion

In developing the checklist, we identified supporting scientific evidence for only nine of 18 items to guide the reporting of meta-analyses of randomized trials. Some of this evidence is indirect. For example, under the "Abstract" heading we ask authors to use a structured abstract format. The supporting evidence for this item was collected by examining abstracts of original reports of individual studies and may not pertain specifically to the reporting of meta-analyses. However, the QUOROM group considered this a reasonable approach by analogy with other types of research reports and pending further evidence regarding the merits of structured abstracts for meta-analyses.

We have asked authors (under the "Methods" heading and "Validity assessment" subheading) to be explicit in reporting the criteria used when assessing the "quality" of the trials included in their meta-analysis and the outcome of the quality assessment. Direct and compelling evidence does exist to support recommendations about reporting on the quality of RCTs included in a meta-analysis. A meta-analytic database of 255 obstetric RCTs provided evidence that randomized trials with inadequate reporting of allocation concealment (i.e., keeping the intervention assignments hidden from all participants in the trial until the point of allocation) as compared to those trials in which this information was adequately reported resulted in an overestimation of the intervention effect by 30%. Similar results across several disease categories and methods of quality assessment have been reported recently. These findings suggest the likelihood that inclusion of reports of low-quality RCTs in meta-analyses may alter the summary measures of the intervention effect.

We have asked authors (under the "Methods" heading and "Quantitative data synthesis" subheading) to be explicit in reporting assessment of publication
bias. Likewise, (under the "Discussion" heading) we have recommended comments about whether the results obtained may have been influenced by publication bias. Publication bias derives from the selective publishing of studies with statistically significant or directionally positive results and it can lead to inflated estimates of efficacy in the meta-analyses. For example, trials of single alkylating agents versus multiple agent cytotoxic chemotherapy in the treatment of ovarian cancer have been analyzed. Published trials yielded significant results in favour of the multiple-agent therapy, but that finding was not supported when the results of all trials—both those published and those registered but not published—were analyzed.

We have asked authors (under the "Methods" heading and "Searching" subheading) to be explicit about the publication status of reports included in a meta-analysis. Only about one third of published meta-analyses report the inclusion of unpublished data. Whereas one study found that there were no substantial differences in the dimensions of study quality between published and unpublished clinical research, another study suggested that intervention effects reported in journals were 33% greater than those reported in doctoral dissertations. The role of the grey literature (i.e., literature that is difficult to locate and/or retrieve) was examined in 39 meta-analyses that included 439 RCTs, 120 of which were grey literature. Meta-analyses limited to published trials, compared with those that included both published and grey literature, overestimated the treatment effect by an average of 12%. There is still debate between editors and investigators concerning the importance of including unpublished data in a meta-analysis.

We have asked authors (under the "Methods" heading, and "Searching" subheading) to be explicit in reporting whether they have used any restrictions on language of publication. Approximately one third of published meta-analyses have some language restrictions as part of the eligibility criteria for including individual trials. It is not clear why, since there is no evidence to support differences in study quality, and there is evidence supporting that language restrictions may result in a biased summary. The reports of 127 RCTs written in English, as compared with those reported in four other languages, demonstrated minimal or no difference in several important methodologic features. Similar results have been reported elsewhere. The role of language restrictions has been studied in 211 RCTs included in 18 meta-analyses in which trials published in languages other than English were included in the quantitative summary. Language-restricted meta-analyses, as compared with language-inclusive ones, overestimated the treatment effect by only 2% on average. However, the language-inclusive meta-analyses were more precise.
Reports of RCTs with statistically positive results are more likely to be published in English.\textsuperscript{31} Likewise, there is emerging evidence to suggest that reports of RCTs from certain countries usually have statistically positive results.\textsuperscript{46}

We used several methods to generate the checklist and flow diagram: a systematic review of the reporting of meta-analyses; focus groups of the steering committee; and a modified Delphi approach during the consensus conference. Although we did not involve certain consumers of meta-analyses (e.g., policymakers or patients) we formally pre-tested this document with representatives of several constituencies who would use the recommendations, after which we made modifications accordingly.

The QUOROM group also discussed issues regarding the format of a meta-analysis report, how best to evaluate the impact of the QUOROM statement, and how best to disseminate it. The format we recommend includes nine new subheadings that reflect the sequential stages in its conduct, within the text of the report of a meta-analysis. The checklist included in the statement can also be used during the planning, performing, and reporting of a meta-analysis and during peer review of the report after its submission to a journal.

We delayed publication of the QUOROM statement to first evaluate its impact on the editorial process. We conducted a randomized trial, involving eight medical journals, evaluating the impact of applying QUOROM criteria on journal peer review. Accrual is now complete and we will report the trial results shortly.

After approximately five weeks of electronic posting we had received five comments from investigators whom we thank for their thoughtful consideration of the QUOROM statement. Several issues, in particular regarding the use of terminology, cannot be addressed in the statement at present. The QUOROM group is agreed on the importance of making changes to the checklist in the light of documented evidence and must resist changes based on opinion or anecdotal evidence unless there is a compelling rationale for doing otherwise. Nonetheless, the issues raised have been noted for consideration and discussion in future deliberations of the QUOROM group.

Several queries addressed the distinction between the meta-analysis and systematic review. As is indicated in the introduction, and is now consistent throughout the statement, the QUOROM group agreed to observe the distinction as defined by the Potsdam consultation on meta-analysis.\textsuperscript{3}

We were also asked to clarify the checklist item asking investigators, under the “Discussion” heading, to interpret their results in light of the totality of evidence. Increasingly, several meta-analyses on the same topic are reported.\textsuperscript{47-49} If other similar reports are available, authors should discuss their results as it relates to such evidence.
The QUOROM statement

For the QUOROM statement to continue to be useful it must remain evidence based and up-to date. The QUOROM group needs to survey the literature continually to help inform themselves about emerging evidence regarding meta-analysis reporting. This information needs to be collated and presented annually for two purposes. One, to decide which checklist items to keep, delete, and/or add. These decisions can be made in a manner similar to the selection of the original items. Two, to deliver a state of the union address concerning meta-analysis reporting. All of these efforts being coordinated through a web site. This approach is similar to the CONSORT initiative.

In summary, our choice of items to include in a meta-analysis report was based on evidence whenever possible, which implies the need to include items that can systematically influence estimates of treatment effects. Currently, we lack a detailed understanding of all the factors leading to bias in the result of a meta-analysis. Clearly, research is required to help improve the quality of reporting of meta-analyses. Such evidence may also act as a catalyst for improving the methods by which meta-analyses are conducted.

The QUOROM checklist and flow diagram is available on this journal’s web site (www.thelancet.com). We hope that this document will generate further interest in the field of meta-analysis and that, like the CONSORT initiative, the QUOROM statement will become available in different languages and locations as it is disseminated. We invite interested readers, reviewers and researchers and editors to use the QUOROM statement and generate ideas for improvement.
References


27. Taddio A, Pain T, Fassos FF, Boon H, Illersich AL, Einarson TR: Quality of nonstructured and structured abstracts of original research articles in the
Chapter 10


