

COMMENT

## Assessing the scientific quality of review articles

Ruairidh Milne, Larry Chambers

### Why review articles?

Review articles are useful to those involved in health care, whether they are providers, purchasers, or researchers, for two reasons. Firstly, they are one way of coping with huge volumes of published reports (the number of biomedical journals worldwide is increasing exponentially, with a doubling time of about 19 years<sup>1</sup>). Secondly, impressive examples now exist of the power of systematic reviews to provide reliable answers to important questions where previously the answers were lacking or wrong:<sup>2</sup> for example, the effects of systemic treatment on early breast cancer,<sup>3</sup> or of corticosteroids given to women expected to deliver prematurely.<sup>4</sup>

Until recently, however, a double standard has operated. Whereas authors, journal editors, and journal readers have expected a structured approach to reports of original research (supplemented increasingly by structured abstracts), the same has not been true of review articles. Although widely used to draw conclusions about research, reviews have not been seen as an extension of scientific reporting<sup>5</sup> and little effort has been put into ensuring their quality.

### Problems with the scientific quality of review articles

Empirical evidence of problems with review articles is not lacking. Mulrow identified 50 review articles published in 1985–6 and assessed how they performed against eight criteria (table I).<sup>6</sup> Only one review satisfied as many as six of the criteria. Antman *et al* looked at the recommendations of clinical experts (in review articles and textbook chapters) in the acute treatment and secondary prevention of myocardial infarction.<sup>7</sup> Using the technique of “cumulative meta-analysis”, they used evidence from randomised controlled trials to calculate retrospectively the apparent odds ratios of benefit associated with particular treatments in particular years. They identified important discrepancies between the experts’ recommendations and what might have been known from systematic reviews and concluded that clinical experts needed access to better databases and new statistical techniques. Another conclusion must be that readers would make more accurate inferences if they could assess reviews’ scientific quality.

### What is “a review”?

Journal editors have been reluctant to introduce structured abstracts for review articles, even

though this would make it easier for readers to assess them. One reason is the sense that this is overkill. One would not, surely, expect readers to assess book reviews for their scientific quality? Part of the problem lies with the word “review” itself, which has many meanings. One suggested by the *Concise Oxford Dictionary* is relevant here: “general survey or reconsideration of subject or thing”.<sup>8</sup> But health care reviews of this sort are not homogeneous; they span a spectrum, from those that describe a wide range of topics to those that attempt to answer a precise question. This article, for instance, is an unsystematic review that advances an argument—that the scientific quality of review articles needs assessing—rather than one that tries to synthesise evidence in a systematic way. Does the same straitjacket need to be imposed on all reviews?

No: but readers do need to be able to assess the scientific quality, and so the reliability and usefulness, of any review article, whether it is called a review, an editorial, or a consensus statement. They can be helped in this by structured abstracts and checklists (for example tables I and II). Just as important is improving the quality of reviews themselves, something that will be helped by the development of the Cochrane Collaboration. This has been catalysed by the establishment of the UK Cochrane Centre in 1992,<sup>2</sup> the main purpose of

Table I Criteria for assessing review articles<sup>6</sup>

1	Was the specific purpose of the review stated?
2	Were sources and methods of the citation search identified?
3	Were explicit guidelines provided that determined the material included in and excluded from the review?
4	Was a methodologic validity assessment of material in the review performed?
5	Was the information systematically integrated with explication of data limitations and inconsistencies?
6	Was the information integrated and weighted or pooled metrically?
7	Was a summary of pertinent findings provided?
8	Were specific directives for new research initiatives proposed?

Table II Guidelines for reading reviews of published reports<sup>10</sup>

1	Were the questions and methods clearly stated?
2	Were comprehensive search methods used to locate relevant studies?
3	Were explicit methods used to determine which methods to include in the review?
4	Was the validity of the primary studies assessed?
5	Was the assessment of the primary studies reproducible and free from bias?
6	Was variation in the findings of the relevant studies analysed?
7	Were the findings of the primary studies combined appropriately?
8	Were the reviewers’ conclusions supported by the evidence cited?

Directorate of Health Policy and Public Health, Oxford Regional Health Authority, Headington, Oxford  
R Milne  
Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada  
L Chambers

Accepted for publication March 1993

which is to facilitate systematic, up to date reviews of randomised controlled trials of health care, such as those already conducted in the fields of pregnancy and childbirth.<sup>9</sup> Together, these developments—of higher quality reviews and of readers better able to assess their quality—suggest that reviews in the future will have a major impact on knowledge, practice, and policy in health services.

- 1 Wyatt J. Uses and sources of medical knowledge. *Lancet* 1991; **338**: 1368–73.
- 2 Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992; **305**: 786–8.
- 3 Early breast cancer trialists' collaborative group. Systemic treatment of early breast cancer by hormonal, cytotoxic, or immune therapy (part 1). *Lancet* 1992; **339**: 1–15.
- 4 Crowley P, Chalmers I, Keirse MJNC. The effects of corticosteroid administration before preterm delivery: an overview. *Br J Obstet Gynaecol* 1990; **97**: 11–25.
- 5 Dickersin K, Berlin JA. Meta-analysis: state of the science. *Epidemiol Rev* 1992; **14**: 154–76.
- 6 Murlow CD. The medical review article: state of the science. *Ann Intern Med* 1987; **106**: 485–8.
- 7 Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomised control trials and recommendations of experts. *JAMA* 1992; **268**: 240–8.
- 8 *Concise Oxford Dictionary*. 7th edition. Oxford: Oxford University Press, 1982.
- 9 Chalmers I, Enkin M, Keirse MJNC (eds). *Effective care in pregnancy and childbirth*. Oxford: Oxford University Press, 1989.
- 10 Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *Can Med Assoc J* 1988; **138**: 697–703.