

Guyatt GH, Oxman AD (2009). Medicine's methodological debt to the social sciences.



© Gordon Guyatt, McMaster University Medical Centre, Room 2C12, 1200 Main Street West, Hamilton, Ontario L8N 3Z5, Canada. Email: guyatt@McMaster.ca

The editor of the James Lind Library, Iain Chalmers, invited us to document the origins of a 1988 paper, published in the *Canadian Medical Association Journal*, in which we drew attention to the need to assess the methodological quality of medical review articles ([Oxman and Guyatt 1988](#)). Our paper was one of two such papers published in the general medical journals at that time, the other having been published a few months earlier by Cynthia Mulrow ([Mulrow 1987](#); [Huth 2008](#)).

Our 1988 article was based on work done by one of us who was a student at the time, supervised by the other, who was a faculty member, for a Master's degree at McMaster University, Canada (Oxman 1987). Iain Chalmers suggested the interview format which follows, in which **GHG**, the faculty member, interviews **ADO**, the former student.

Cite as:

Guyatt GH, Oxman AD (2009). Medicine's methodological debt to the social sciences. JLL Bulletin: Commentaries on the history of treatment evaluation (www.jameslindlibrary.org). [Personal reflection]

GHG: You arrived at McMaster University in 1984. Had you thought about systematic reviews before arriving?

ADO: No

GHG: When did the notion of the importance of the scientific quality of reviews first occur to you? What put the idea into your head?

ADO: Both as a general practitioner, prior to becoming a resident in community medicine, and as a student, I had been confronted with the shortcomings of traditional reviews. As a general practitioner in a remote community, like most busy clinicians, I did not have the luxury of reviewing the primary research myself for most of the clinical problems with which I was faced in practice. I relied extensively on reviews published in textbooks and in journals. For controversial questions, including how to manage common problems such as hypertension or otitis media, my unsatisfactory solution was to rely on what I knew about the prestige of the author, the institution, or the journal to decide what was right. Of course, this did not always lead to correct conclusions, because of my limits in judging prestige and because often people of equal prestige disagree. Besides which, there is no reason to believe that prestigious people are any less biased or more scientific in how they integrate evidence than the rest of us.

When I had to choose a thesis topic after going to McMaster in 1986, my choice grew out of a reading course in which I investigated the role of evidence in public health decision-making, an area of obvious interest to me as a resident in community medicine. In the Design, Measurement, and Evaluation (DME) MSc programme at McMaster, the superiority of the randomized controlled trial as a study design is heavily emphasized. Yet in public health, preventive medicine, occupational health and environmental health, randomized controlled trials are hard to find and are often impossible to undertake. There is often a need to integrate evidence from a variety of study designs, and with varying degrees of relevance to a given question.

When I began to read about how to integrate the evidence relevant to any particular question, I was inspired by a book that I came across in the University book store, written by two social scientists, Richard Light and David Pillemer. In the book they had eloquently summarized the state of the art of the "science of reviewing research" ([Light and Pillemer 1984](#)). Subsequently, Greg Jackson's earlier investigation of how social scientists review research influenced me (Jackson 1978). Within medicine, we owe a lot to these and other social scientists that pioneered thinking about and the development of methods for research synthesis.

GHG: I think you were inclined initially to call what we now call systematic reviews "overviews". Do you have any reflections on that evolution?

ADO: The terms 'research overview', 'research review', 'research synthesis', 'integrative review', and 'review' have

tended to be used fairly synonymously to refer to the summarization and synthesis (integration of independent research studies on a given topic), and have included both qualitative approaches (literature reviews, narrative reviews, or traditional reviews) and quantitative (statistical) methods ('meta-analyses').

As best I can recall, my decision to use the term 'research overviews' in the mid-1980s was made in response to some of the answers to a survey of medical journal editors about the criteria they used for evaluating reviews. Several editors were confused by what we meant by 'research reviews', which is the term we used in the letter we sent them (Oxman 1987, p 215). In addition to revealing a problem with the terminology, the survey laid bare a lack of standards and methodological criteria for assessing research overviews. Most editors reported relying on experts rather than methods. For example, one said "we rely heavily on the expertise of our individual Committee members as guided by the group to choose qualified reviewers. We thus monitor the quality of the manuscripts before they are even written!" Only one editor at the time reported requiring the use of scientific methods and for those to be spelled out in review articles. He noted that "As a result, we publish very few review articles."

Some other articles and books contemporary to our 1988 article were using the terms 'overview' (for example, Yusuf et al. 1985; Early Breast Cancer Trialists' Collaborative Group 1988) and 'meta-analysis' (for example, L'Abbé et al. 1987; Sacks et al. 1987; [Jenicek 1987](#)). And some writers and readers tended to concentrate on statistical synthesis rather than the methods needed to reduce biases in reviews. In their foreword to a book that they had deliberately entitled *Systematic Reviews*, Chalmers and Altman made a plea for separate methodological challenges (minimizing bias, and reducing the play of chance) to be distinguished by reserving the term 'meta-analysis' for the process of statistical synthesis to reduce the play of chance ([Chalmers and Altman 1995](#)). This view was subsequently reflected in John Last's *Dictionary of Epidemiology* (Last 2001). The term 'systematic reviews' has now been very widely adopted, but confusion continues to exist in some quarters.

GHG: When did you start and complete your thesis, and publish material based on it?

ADO: I started researching and writing the thesis in 1986, and defended and 'published' it in 1987. Based on the work, we drafted guidelines for reading literature reviews and submitted a paper proposing these to the *Annals of Internal Medicine*, at the time that Cindy Mulrow's article came out ([Mulrow 1987](#); [Huth 2008](#)). The work based on the protocol in my MSc thesis appears to have been an early use of a systematic approach to developing and evaluating critical appraisal criteria as a type of measurement instrument.

GHG: Presumably your findings prompted you to develop a systematic reviews training course at McMaster. When did you think of creating the course? Can you describe its evolution?

ADO: Ideas for the course emerged while working on my thesis, and plans for it were developed while I was writing up the work. I think we first offered the course in the fall of 1987. In my thesis I reviewed the methodological literature for each step in undertaking a review. That formed the basis for the course, which was developed from my thesis. Students were expected to have a question at the beginning of the course and to complete a review by the end of the course. Each week we would work through a module with background reading, examples, and exercises where the students would apply what they were learning to their own reviews.

GHG: What was the contribution of the course to the evolution of your ideas?

ADO: We had Tom Chalmers and Larry Hedges as guest 'lecturers' on the course, and I expect we learned as much as the students did the first times we offered it (Iain Chalmers and Murray Enkin were students on the first course). Several of the reviews done for the course were published, and many of those who took the course became active in the Cochrane Collaboration. The course was challenging for students and proved to be a great way to pull together and consolidate what they had learned from the Design, Measurement, and Evaluation (DME) Masters of Science programme at McMaster.

When we started there were not a lot of published systematic reviews in healthcare. Research synthesis was an emerging science and our ideas and knowledge evolved with the course. A lot of new methodological articles came out that became part of the course material. Discussing review methods with (very bright) students and helping them to apply the methods to a wide range of questions helped to clarify our thinking (for example, about subgroup analyses), as well as to identify and clarify new methodological challenges (for example, systematic reviews of diagnostic test accuracy).

These new methodological challenges have certainly been intellectually challenging for me. While at McMaster, I helped to apply systematic review methods to environmental health questions. Later, together with Brian Haynes, Dave Davis, Jeremy Grimshaw and other contributors to the Cochrane Effective Practice and Organisation of Care Group, and others, I have been involved in developing methods for addressing complex questions and for applying these to questions about improving practice, health systems and health policy, in low and middle-income countries, as

well as richer countries.

GHG: What do you see as your role in the development of systematic review methodology?

ADO: This is really for others to judge. I think it has become less likely that clinicians will find it as hard as I did when I was in practice to access reliable evidence to inform their decisions. My MSc thesis provided me with an opportunity to review systematically what was known about systematic review methods, and a number of things flowed from this starting point:

- We (you and I) helped to bring attention to work from the social sciences and to make it more relevant to healthcare.
- We helped to bring attention to the shortcomings of relying on experts to synthesize research (Oxman and Guyatt 1993), and we popularized the concept of and need for systematic reviews, for example, in the first *Canadian Medical Association Journal* readers' guide (Oxman and Guyatt 1988), the *BMJ* checklist (Oxman 1994), and the JAMA Users' Guide to the Medical Literature, in which we focused particularly on the importance of formulating answerable questions (Oxman et al. 1993).
- We helped to bring attention to the need for systematic reviews and their role in health technology assessment and clinical practice guidelines.
- We developed teaching materials and applied small-group, problem-based approaches to teaching systematic review methods. This work contributed to subsequent training efforts by the Cochrane Collaboration and formed the basis for the first Cochrane Handbook. It was later developed for raising awareness among journalists and the general public of the need to understand some of the basic concepts of review methods.
- Our guidelines for subgroup analyses (Oxman and Guyatt 1992) helped to bring attention to and provide a structured approach to a common problem in systematic reviews, and in research more generally.
- Most recently, we have tried to make the results of systematic reviews more useful to people making decisions by developing 'Summary of Findings' tables (Schünemann et al. 2008), and, as systematic reviews have become more numerous, we have helped to develop methods for overviews of systematic reviews (Becker and Oxman 2008).

None of these contributions were unique or done alone. They are all part of a large collaborative effort. To reiterate an observation made by John Ziman in an article published in *Nature* many years ago:

"Our present system of rewards and incentives in science does not encourage individuals to devote themselves for years on end to these critical synthesizing activities. 'Recognition', by way of professional advancement and prestige, is given solely for primary research; has any academy, ever mentioned that the hero was the author of a valuable treatise or of the authoritative review that has since determined the course of research in his field?"

"The trouble is, quite simply, a matter of philosophy. We are so obsessed with the notions of discovery and individual originality that we fail to realize that scientific research is essentially a corporate activity, in which the community achieves far more than the sum of the efforts of its members" (Ziman 1969).

Finally, although motivated by a belief that this work is important for the welfare of patients and the public, as a leading member of the Writing Group of 'Clinicians for the Restoration of Autonomous Practice' ([CRAP 2002](#)) I feel it is important not to take oneself too seriously.

References

Becker L, Oxman AD (2008). Overviews of reviews. Chapter 22. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester (UK): Wiley-Blackwell.

Chalmers I, Altman DG (1995). *Systematic Reviews*. London: BMJ Publications.

Clinicians for the Restoration of Autonomous Practice (CRAP) Writing Group (2002). EBM: unmasking the ugly truth. *BMJ* 325:1496-1498.

Early Breast Cancer Trialists' Collaborative Group (1988). Effects of adjuvant tamoxifen and of cytotoxic therapy on mortality in early breast cancer. An overview of 61 randomized trials among 28,896 women. *N Engl J Med* 319:1681-92.

Huth EJ (2008). The move toward setting standards for the content of medical review articles. The James Lind Library (www.jameslindlibrary.org).

Jackson GB (1978). Methods for Reviewing and Integrating Research in the Social Sciences. Final report to the National Science Foundation for Grant no. DIS 76-20309, Social Research Group, George Washington University, Washington, D.C.

Jenicek M (1987). Méta-analyse en médecine. Évaluation et synthèse de l'information clinique et épidémiologique. St. Hyacinthe and Paris: EDISEM and Maloine Éditeurs.

L'Abbé KA, Detsky AS, O'Rourke K (1987). Meta-analysis in clinical research. *Ann Int Med* 107:224-232.

Last JM (2001). *A Dictionary of Epidemiology*. Oxford: Oxford University Press.

Light RG, Pillemer DB (1984). *Summing Up: The Science of Reviewing Research*. Harvard University Press, Cambridge, MA.

Mulrow CD (1987). The medical review article: state of the science. *Ann Intern Med* 106:485-488.

Oxman AD (1987). *A Methodological Framework for Research Overviews*. MSc Thesis, McMaster University, Hamilton, Ontario, Canada.

Oxman AD (1994). Checklists for review articles. *BMJ* 309:648-651.

Oxman AD, Guyatt GH (1988). Guidelines for reading literature reviews. *Can Med Assoc J* 138:697-703.

Oxman AD, Guyatt GH (1992). A consumer's guide to subgroup analyses. *Ann Intern Med* 116:78-84.

Oxman AD, Guyatt GH (1993). The science of reviewing research. *Annals of the New York Academy of Science* 703:125-134.

Oxman AD, Sackett DL, Guyatt GH, for the Evidence-Based Medicine Working Group (1993). Users' guides to the medical literature, I. how to get started. *JAMA* 270:2093-2095.

Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC (1987). Meta-analysis of randomized controlled trials. *New England Journal of Medicine* 316:450-455.

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Glasziou P, Guyatt GH (2008). Presenting results and 'summary of findings' tables. Chapter 11. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester (UK): Wiley-Blackwell.

Yusuf S, Peto R, Lewis J, Collins R, Sleight P (1985). Beta blockade during and after myocardial infarction: an overview of the randomized trials. *Progress in Cardiovascular Disease* 27:335-371.

Ziman JM (1969). Information, communication, knowledge. *Nature* 224:318-324.

[Home](#)

[Contents](#)