Downloaded from http://ebmh.bmj.com/ on October 30, 2017 - Published by group.bmj.com

Understanding and interpreting systematic reviews and meta-analyses. Part 1: rationale, search strategy, and describing results

Reviews of primary studies are an important source of information for clinicians, and we include abstracts of good quality reviews in *Evidence-Based Mental Health*. This issue contains summaries of reviews of the effectiveness of psychotherapy (Wampold p 78), cognitive therapy in depression (Gloaguen p 76), family and couples therapy for drug abuse (Stanton p 81), chlorpromazine in schizophrenia (Thornley p 83), interventions for old age depression (McCusker p 77), support after post-partum depression (Ray p 89), relative mortality in schizophrenia (Brown p 91) and the association between apolipoprotein E and Alzheimer's disease (Farrer p 94). All of these reviews tried to access and review systematically all of the relevant articles in the field (**systematic review**) and included a quantitative summary of their results (**meta-analysis).** The clinical interpretation of a review article, or an abstract of a review article, requires the reader to have some conceptual understanding of how systematic reviews are conducted and the rationale behind the approach. We address these issues in this article and discuss two important stages of a review—setting a clear research question and identifying the primary studies. We will focus on systematic reviews of treatment studies although similar principles apply to reviews of different kinds of studies.

In part 2 (November issue), we will describe some of the statistical issues that need to be considered when interpreting the results of a systematic review that uses statistical techniques to combine data from different studies (**meta-analysis**).

The need for systematic reviews

One of the more important methodological "discoveries" of the past two decades was that many review articles of health interventions were methodologically inadequate.[1] Although there had been significant advances in research design, such as the development of the randomised controlled trial, the review article still tended to be unsystematic and susceptible to many biases. The result was that it was difficult to assess how unbiased the reviewer was in reaching his or her conclusions. Often, the author of a review article had a particular viewpoint and only included studies that supported this position. There was a need for *systematic reviews* in which the author gave full details of the methods used to identify the primary studies. The methodology of reviewing research had been developing gradually for several decades. One landmark contribution was from within the field of mental health, the pioneering systematic review of psychotherapy outcome studies that introduced the term *meta-analysis*.[2] Further developments led to the establishment of the Cochrane Collaboration in 1993, an international organisation dedicated to producing regularly updated systematic reviews covering all healthcare interventions.[3] However, poor quality reviews still abound and clinicians need to know how to tell a valid review from an invalid one. If a meta-analysis (this term now tends to be used to describe the statistical summary of results within a systematic review) has been performed, clinicians need to know how to tell if it was justified and how to interpret the results. To try and help, we only include abstracts of systematic reviews in *Evidence-Based Mental Health* if they meet at least our minimal methodological criteria (see purpose and procedure). Our aim is to produce rigorous summaries of reviews of high quality published research that are relevant to clinical practice.

How to appraise critically a review paper

The main feature that distinguishes a systematic review from an unsystematic review is a methods section that adequately describes the research question, the search strategy, and the designs of the studies that were selected. By reading the methods section, the user of the review can decide how valid, and in particular, how free from bias, the reviewer's conclusions are likely to be.

DID THE REVIEW SET OUT TO ANSWER A CLEARLY DEFINED QUESTION?

The first thing to assess is whether the review was adequately focused on a particular clinical question. The nature of this clinical question (ie, does it concern diagnosis, treatment, prognosis, economic evaluation, etc) will define the research designs that the reviewer should include.

An important consideration in high quality research is the development of a research plan or protocol. The same is true for systematic reviews. Careful consideration of inclusion and exclusion criteria for studies, methodological issues to be assessed, and planned comparisons is likely to be a good investment of time and lead to greater efficiency in the conduct of the review. There may subsequently be good reasons to depart from this plan to some degree—especially if the protocol proves impractical in the light of the available studies. Occasionally, a relevant comparison becomes apparent during the conduct of a review and, ideally, an article describing a systematic review will indicate when this is the case. For the reader, knowledge that a protocol driven research design is used provides increased confidence that reviews were not unduly influenced by the inclusion or exclusion of specific studies or by the selection of specific outcomes by the results observed. Clearly, including only the positive studies that examine the effects of an intervention, and excluding all the negative ones, will lead to an overly optimistic estimate of the benefits of treatment.

WHAT SEARCH STRATEGY WAS USED TO IDENTIFY THE PRIMARY STUDIES?

The lengths to which reviewers go to identify relevant studies can affect their conclusions. Whereas patients are the subjects

Downloaded from http://ebmh.bmj.com/ on October 30, 2017 - Published by group.bmj.com

of many clinical trials, trials may be considered the subjects of systematic reviews. Thus it is important that reviews reflect all relevant trials rather than a subset of them. This is partly because gaining additional trials will often add extra useful information and will increase the accuracy and precision of estimates of the effects of treatment. At least as important is the risk of publication bias. Trials with positive and interesting results are often easier to access than those that do not have statistically significant findings (**negative studies**). In practice, negative studies may be more likely to remain unpublished. Thus the reader will want to be reassured that reviewers have gone to reasonable lengths to access all studies that may fall within the inclusion criteria of the review. In that way, the risk of summarising the results of a only a subset of relevant studies with the most promising results, can be avoided.

Many reviewers use electronic databases such as Medline, EMBASE, or PsycLIT for their searches—although it may be that as few as 50% of relevant articles will be identified by this method.[4] In this issue of *Evidence-Based Mental Health*, Brown (p 91) used Medline and BIDS (an UK academic database). Farrer *et al* (p 94) used Medline, and Wampold *et al* searched "journals which typically publish studies comparing ≥2 bona fide psychotherapies." All of these authors used a systematic, reproducible approach to identifying the primary studies. A Cochrane review we have abstracted in this issue (Thornley *et al*, p 83) used an optimally sensitive and comprehensive search strategy including electronic databases and hand searching of journals. The reviewers also attempted to identify unpublished studies by contacting the drug company who made chlorpromazine. This may be considered a gold standard search.

DID THE SELECTED STUDIES USE THE APPROPRIATE RESEARCH DESIGN AND HOW WAS THE QUALITY OF THE PRIMARY STUDIES ASSESSED?

In *Evidence-Based Mental Health* we abstract systematic reviews of studies on the full range of clinical topics. The most appropriate design of primary study will depend on the type of research question. Many systematic reviews are of treatment studies and we will focus on these in this article and in part 2 (November issue). Our minimal quality criteria for selecting reviews are described in the purpose and procedure (p 66). One of these criteria for treatment studies is that at least some of the included studies are randomised controlled trials. We have previously described the advantages of the randomised controlled trial—when treatment effects are moderate and additional factors which may affect treatment outcome are relatively large, random allocation is the best way available for avoiding bias in the distribution of known and unknown biases (see the EBMH notebook by Peter Szatmari, issue 2, titled "some useful concepts and terms used in articles about treatment"). Random allocation also provides a valid basis for an estimate of the magnitude of bias and thus for the width of confidence intervals around an estimate of the effect of an intervention.

Ultimately, the quality of the review and the degree of faith the reader can have in the results will depend on the quality of individual trials included. The systematic review should therefore give some indication of the quality of the individual trials and how this was assessed. Essentially, randomised controlled trials are simple things—but there are several biases that can lead to invalid results. The main points to check are that there was adequate concealment of allocation, that the rate of drop out was minimised, and that the patients were analysed in the groups to which they were allocated. Obviously, the study must also use a robust, clinically relevant measure of outcome that was objectively assessed. Checklists for assessing the quality of randomised controlled trials are widely available.[5] From the point of view of the user of a review article, the important thing is to make sure that the authors have used a reasonable and explicit method of assessing the primary studies. The authors should ideally have had at least 2 people evaluating the quality of a subset of the studies, and reported the agreement between the raters.

Comment

In *Evidence-Based Mental Health* we attempt to describe the major features in the design of abstracted studies that may affect their quality. We also attempt a common format in the descriptions of the results to aid readers in the task of interpreting study findings in the light of their existing practice. In the next article we will describe some of the major features of meta-analysis (the combination of the results from different studies) which are commonly used in systematic reviews that we abstract, but which remain the subject of some controversy.

JOHN GEDDES, MD
Editor, Evidence-Based Mental Health

NICK FREEMANTLE, MA
Medicines Evaluation Group, Centre for Health Economics, University of York, UK

DAVID STREINER, PhD
Editor, Evidence-Based Mental Health

SHIRLEY REYNOLDS, MSc
Editor, Evidence-Based Mental Health

1 Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987;**106**:485–8.
2 Smith ML, Glass GV. Meta-analysis of psychotherapy outcome studies. *American Psychologist* 1977;**32**:752–60.
3 Fullerton-Smith I. How members of the Cochrane Collaboration prepare and maintain systematic reviews of the effects of health care [EBM note]. *Evidence-Based Medicine*. 1995 Nov-Dec;**1**:7.
4 Adams CE, Power A, Frederick K, *et al*. An investigation of the adequacy of Medline searches for randomized controlled trials (RCTs) of the effects of mental health care. *Psychol Med* 1994;**24**:741–8.
5 Sackett DL, Richardson S, Rosenberg W, *et al*. *Evidence-based medicine. How to practise and teach EBM*. London: Churchill-Livingstone, 1997.