

HOW TO DO A SYSTEMATIC LITERATURE REVIEW

ABSTRACT: *Physiotherapists, whether serving individual patients or populations, always have to sought to base their decisions and actions on the best possible evidence. In making choices, health professionals may benefit from structured summaries of the options and outcomes, systematic reviews of the evidence and recommendations regarding the best choices. The aim of this paper is to present guidelines on how to conduct a systematic review. The structure and content of a systematic review are being discussed, following a step-by-step approach.*

**PAPADOPOULOS M, MSc (Physio)¹
RHEEDER P, (M Med)²**

¹Physiotherapy Department, University of Pretoria
²Clinical epidemiology, University of Pretoria

KEYWORDS: *SYSTEMATIC LITERATURE REVIEW, PHYSIOTHERAPY, METHODOLOGY, META-ANALYSIS*

INTRODUCTION

The problem of how to keep abreast of the literature, both old and new, confronts all health professionals, including physiotherapists. It is clear that most physiotherapists engaged in clinical practice get little or no time for reading during their working day. Because the volume of information available to health care professionals is so large, it is likely that most rely on reviews to deliver synopses of current knowledge in accessible and manageable formats. Reviews are pieces of research in themselves. They use primary research as their data; in other words they are research on research.

Health professionals make many patient management decisions, on a daily basis. Some of these decisions are relatively inconsequential while some others are vitally important. Decisions involve weighing benefits and risks, and institute courses of action judged to be in the patient's best interest. In making choices, health professionals may benefit from structured summaries of the options and outcomes, systematic reviews of the evidence and recommendations regarding the best choices. When properly carried out on the highest possible proportion of all relevant trials, these systematic reviews provide the most accurate and authoritative guides to therapy. Systematic reviews of therapy is such a logical step in progress toward evidence-based health care that it has become the focus of a rapidly growing

international group of clinicians, methodologists and consumers who have formed the Cochrane Collaboration. Systematic reviews that are beginning to flow from this collaboration, and which are updated each time an important new trial is reported, are providing the highest levels of evidence ever achieved on the efficacy of preventative, therapeutic and rehabilitative regimens (Cochrane Database of systematic reviews).

TYPES OF LITERATURE REVIEWS

There are two types of literature reviews: The narrative literature review and the systematic literature review. A narrative literature review typically presents a series of studies, with strengths and weaknesses discussed selectively and informally by one or more acknowledged experts (Rosenfeld, 1996; Oxman et al, 1994). This unsystematic approach of performing a literature review represents the traditional strategy of accumulating and summarising evidence in an unsystematic fashion and then applying implicit preferences to arrive at treatment recommendations. The approach is open to bias and is likely to lead to consistent, valid recommendations only when the gradient between beneficial and adverse consequences of attentive actions is very large. The risk of such unsystematic identification and collection of evidence is that treatment effects may be underestimated or, more commonly, over estimated. Adverse effects may be

exaggerated or ignored. One result of these unsystematic approaches may be recommendations advocating harmful treatments and failing to encourage effective therapy (Guyatt et al, 1999). For example, the observation that patients with ventricular ectopic beats following myocardial infarction were at high risk of sudden death, coupled with the demonstration that these extra beats could be suppressed by specific drugs, formed a rationale for the widespread prescription of these drugs to post-infarction patients with unstable cardiac rhythms. However, subsequent randomized controlled trials examined hard clinical outcomes, not physiologic processes, showed that several of these drugs increase, rather than decrease, the risk of death in such patients and their routine use is now strongly discouraged (Sackett et al, 1997).

A more acceptable alternative would be to collect and read all the available evidence and perform a systematic review. A systematic review is a means of achieving an adequate overview of the literature with respect to the quality of

CORRESPONDENCE:

M Papadopoulos
Department of Physiotherapy
University of Pretoria
PO Box 667
Pretoria
0001
Tel: (012) 354-2023

the reviewed papers and the outcome of the interventions studied (De Bie, 1996).

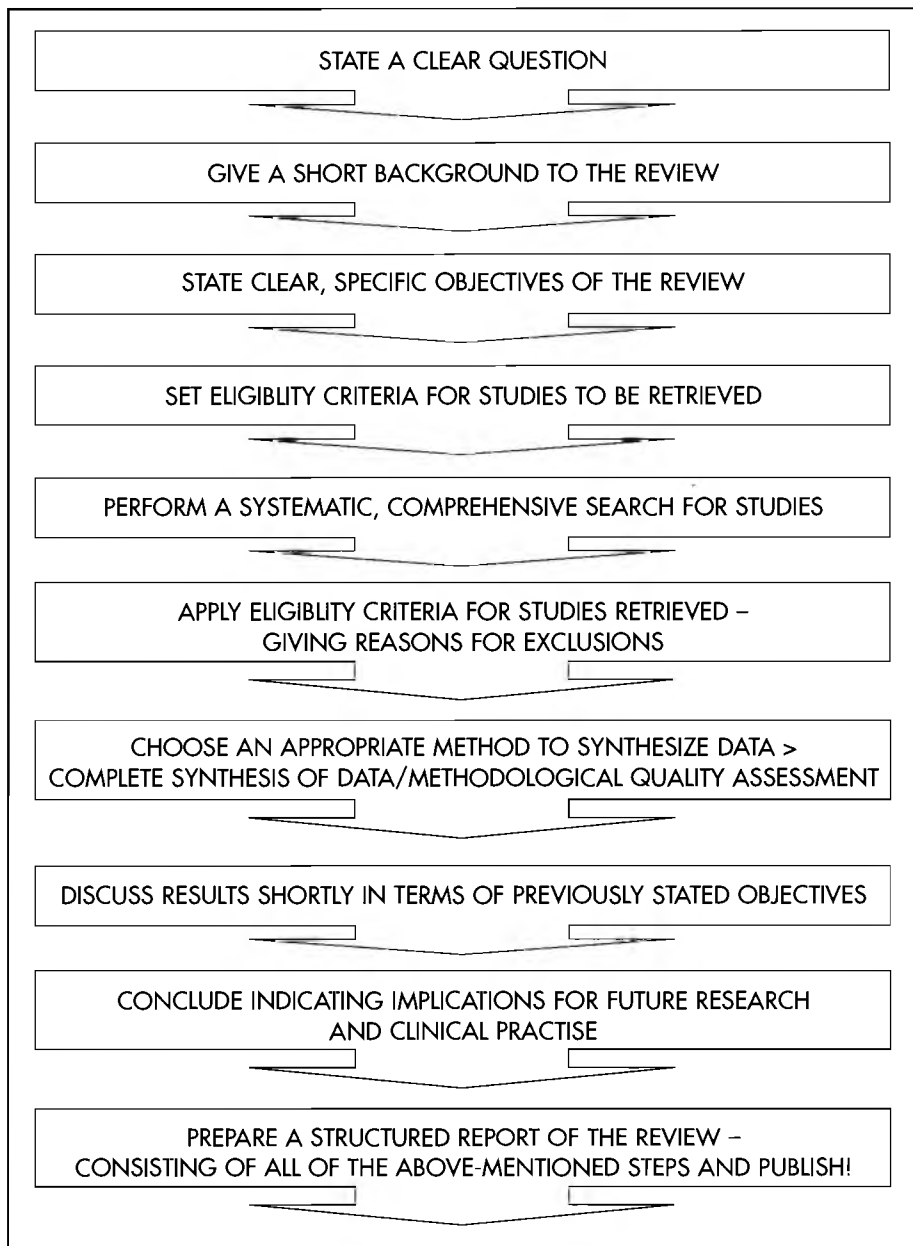
Systematic reviews explicitly state inclusion and exclusion criteria for evidence to be considered, conduct a comprehensive search for the evidence, and summarise the results according to explicit rules (Guyatt et al, 1999). Systematic reviews thus reduce large quantities of research into key findings in a reliable way and offer a means of enabling health care professionals to keep abreast of research; as such they are an essential step in the identification of evidence – based health care (Droogan & Cullum, 1998).

A systematic review provides strong evidence when the quality of the primary studies is high and sample sizes are large and less strong evidence when designs are weaker and sample sizes small. Because judgement is involved in many steps in a systematic review (including specifying inclusion and exclusion criteria, applying these criteria to potentially eligible studies, evaluating the methodological quality of the primary studies, and selecting an approach to data analysis), systematic reviews are not immune from bias. Nevertheless, in their rigorous approach to collecting and summarising data, systematic reviews reduce the likelihood of bias in estimating the causal links between management options and patient outcomes (Guyatt et al, 1999).

AIM OF THIS PAPER

The aim of this paper is to present guidelines on how to conduct a systematic review. To study the quality of trails and summarize the results of well-performed trials with high internal validity, and hence derive valid conclusions, is a rather new phenomenon in physiotherapy. Systematic reviewing could represent a useful tool for not only evaluating the effectiveness of physiotherapy, but also for formulating new hypotheses and identifying areas where lack of knowledge hinders further development (De Bie, 1996). Systematic reviews can also have an educational function in that they could be used for theses or dissertations, thereby indicating where future research can be improved by avoiding flaws from previous research.

TABLE 1: Steps in Performing a Systematic Review



STRUCTURE AND CONTENT OF A SYSTEMATIC REVIEW

A systematic review is composed of a number of sections and steps (see Table 1).

STATES A CLEAR QUESTION

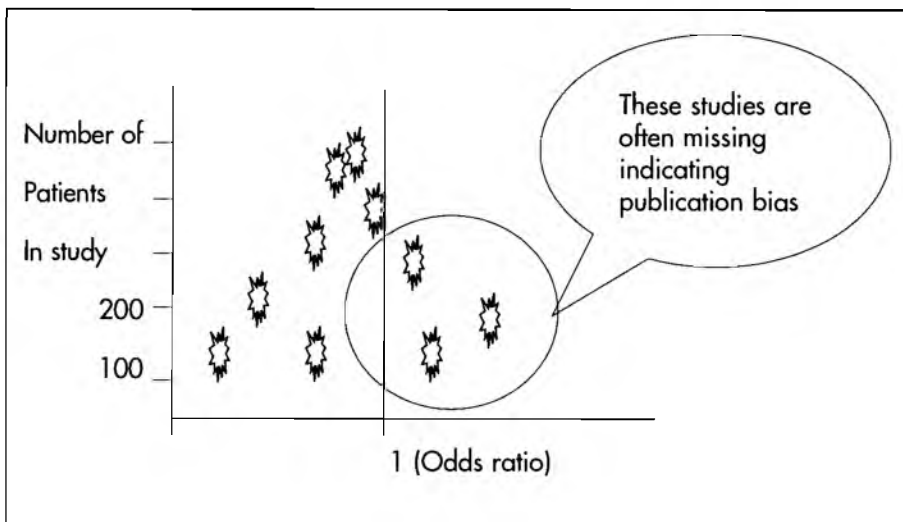
A specific question or set of questions should be stated. The questions should be well-defined and include references to the target population, the health care intervention and the outcomes of interest. A short background to the review should be given stating the importance of the review.

Consequently, the objectives of the review should be stated. Objectives

should be clear, specific and can include sub-group analysis eg whether effects of a certain intervention are dependent on age, gender or the co-existence of other illness or disease etc. For example, to determine the effect of a certain treatment method in patients of different age groups. Will the outcomes in the different age groups be the same?

On the basis of the research question, eligibility criteria are formulated on which studies are to be retrieved. These criteria apply to the participants of studies (exclusion/inclusion criteria), the types of disease or complaint to be studied, the study intervention and the outcome measures to be reviewed.

FIGURE 1: The funnel plot for assessing publication bias. An odds ratio of one represents the line of null effect. Values <1 represent beneficial therapy while values > 1 represent harmful therapy.



Requirements are also set for the study design i.e randomised control trials (RCTs), control trials (CT) or pseudo-randomised control trials (pseudo-RCTs). It should be noted though that the RCT is considered the most reliable method of assessing the efficacy of health care interventions. In the field of physiotherapy research, as in many other disciplines, unfortunately not all trials live up to these standards. Although it remains questionable as to whether evidence should be gathered from less adequately performed studies, it would be foolish to ignore the potential for obtaining valuable information from such investigations (De Bie, 1996).

PERFORMS A SYSTEMATIC, COMPREHENSIVE SEARCH STRATEGY

This is important in order to identify as large as possible a proportion of all the primary studies. A review, which does not make strenuous efforts to identify all relevant studies, runs a great risk of introducing bias. Published material can be found in databases such as the Cochrane library, Medline, Cinahl or Embase. Details of the search such as cut-off dates for studies, key words and MeSH terms should be stated clearly in order to make the search reproducible for other researchers. When reviewing physiotherapy research, a major problem is that a number of trials appear in non-indexed journals. There is also evidence to suggest that studies with “positive”

results are more likely to appear in the published literature (Cook et al, 1993). This phenomenon, called “publication bias”, can yield false positive findings if one or more unpublished or unindexed “negative” articles are excluded from the systematic review. Handsearching of journals should be done as neglecting to do so may result in running the risk of missing important research. Limiting searches to English language articles only, may be a major reason for non-representativeness of studies identified.

Writing to experts, performing library handsearching in order to identify student projects, theses and conference proceedings, are measures required to identify unpublished material. After obtaining all relevant studies, criteria for eligibility, as stated before, should be applied.

The presence of publication bias can easily be detected using a simple graph called a funnel plot (Figure 1). Smaller studies can be expected to yield more variable results and larger studies more precise estimates. If the right side of the funnel only has a small number of studies one can assume that a significant number of negative studies has not been published.

Another statistical method is sometimes used to determine how important missed negative studies would be. This so called “failsafe-n” gives the number of unpublished studies that would be needed to yield the found pooled estimate invalid (Greenhalgh 1997).

APPROPRIATE SYNTHESIS OF DATA

In order to prevent bias a blinded assessment system can be used for the systematic review. Authors, institutes or centres where the studies have been performed and the journal in which the paper was published, can be masked. At least two reviewers should independently study characteristics and outcomes, including information on morbidity and mortality, interventions, length of trial follow-up etc. The final results can be reached by consensus amongst reviewers. Experts from different areas such as statisticians and research methodologists can be used to rank data.

Where there is more than one study sufficiently similar in terms of the intervention studied and the outcome(s) measured, then statistical techniques may be used to synthesis (or pool) data from the primary studies (a process known as meta-analysis). Pooling of data, however is not always appropriate as in studies which are not sufficiently similar in either a clinical or design sense as is very often the case in physiotherapy studies. In such a case a critical review which focuses on the quality of existing studies and the strength of the evidence they provide, will be more appropriate (De Bie, 1996).

Reviewers can use a list of parameters in order to measure the methodological quality of the studies. A weighting scale can be applied to these parameters to discriminate between major and minor flaws in trial design. Parameters that could be considered are listed in table 2. A score of one point (+) can be given to each parameter fulfilled and zero (-) if not. Parameters can be graded as 0,5 points (⊔) if the descriptions are unclear or if incomplete for intervention, outcome measures or data presentation (eg an unclear prescription of blinding can be graded ±). In this way a maximum methodological quality score can be calculated and a hierarchical list, where higher scores indicate studies of higher methodological quality, can be generated.

PRINCIPLES OF QUANTITATIVE SYNTHESIS: META-ANALYSIS

The principle behind meta-analysis is that the larger the sample size of a study the more precise the results. Many

TABLE 2: Suggested Parameters to determine the Methodological Quality of a Study.

- Well described inclusion criteria.
- Presentation of relevant baseline characteristics.
- Random allocation procedure described (allocation concealment).
- Interventions well described (nature, number, duration of treatment).
- Blinding (patients, therapist, observer).
- <10% withdrawals and withdrawals described.
- Effect/outcome measurements relevant and well described.
- Analysis and presentation of the results so that the analysis can be assessed.

statistically insignificant results are found simply because the study was too small (a so called type II error). By pooling different studies (given that the methodologies and outcomes were approximately the same) the sample size is increased and a more precise estimate is obtained (so called pooled estimate, often a pooled odds ratio) (Greenhalgh,1997).

Prior to pooling studies the reviewer must ensure that studies are homogenous (combining apples with apples and not apples with pears). This implies that studies come from the same population and that their results only differ because of chance variation. By visually inspecting the 95% confidence intervals of the estimates one can see whether these overlap or not. If they do not overlap then these studies are heterogenous (implying that they do not come from the same population). This is can also be shown with a statistical test called the chi-square test for heterogeneity with a p-value > 0.05 showing that the studies are indeed homogenous. This test however has low power for detecting heterogeneity (Greenhalgh, 1997).

The statistical method actually used for calculating the "average" or pooled outcome depends on whether heterogeneity was found or not. If they are homogenous the outcome is pooled across studies using a so called fixed-effects model and if they were heterogenous using a random-effects model.

The random-effects model gives a more conservative result (a larger p value). It is also important that heterogeneity be examined further and that results are not just simply pooled. Causes of hetero-

geneity might be related to age, stage of disease or in the case of drugs, dosages used (Greenhalgh 1997).

Often the statistical analysis includes a sensitivity analysis. In doing the effect on the outcome is explored if you manipulate certain variables such as quality weights, inclusion and exclusion of weaker studies etc. The more robust the estimate the better. If however the slightest change in quality weights for instance changes the estimate significantly then there are serious concerns about its validity (Greenhalgh, 1997).

DISCUSSION OF FINDINGS

The most important goal of systematic reviewing is to summarize evidence from original research and thereby provide an adequate overview of the effect or effects of the intervention of interest being studied (De Bie, 1996). In the summary of the systematic review, methodologic limitations of both the primary studies and the review process itself should be elaborated.

This section should thoroughly address the issue of bias in observational studies and the potential role of the placebo effect when a control group is not present (Rosenfeld, 1996). A short discussion of the objectives initially stated, should complete this section.

The conclusions should include a future research agenda including clinical and methodological requirements of proposed studies.

PUBLISHING RESULTS

A properly performed systematic review will qualify for publication in most peer-reviewed journals. The abstract should

be structured, stating clearly that a protocol was followed, that specific objectives were set and that eligibility criteria for article selection were used. A table with key elements of articles such as study design, sample size, treatment dose and duration, outcome measures, patients lost to follow-up etc, is a very helpful tool in supplementing the text. Another table, reporting rating methods and the methodological quality of studies (as in table 2) can also be added. Reasons for excluding identified studies from the systematic review should also be given.

REFERENCES

- De Bie RA; 1996. Methodology of systematic reviews: An introduction. *Physical Therapy Review* 1:47-51.
- Droogan J, Cullum IV; 1998. Systematic reviews in nursing. *International Journal of Nursing Studies* 35:13-22.
- Cochrane database of systematic reviews. Available from BMJ Publishing group, PO Box 295, London WC1H 9TE, UK. Tel: +44 (0) 171 383 6185/6245; Fax +44 (0) 171 383 6662.
- Cook DJ, Guyatt GH, Buan G, Clifton J, Buchingham L, Willan A, McIlroy W, Oxman AD. (1993). Should unpublished data be included in meta-analyses? Current convictions and controversies. *JAMA* 269 (21): 2749-2753.
- Greenhalgh T ; 1997. How to read a paper: Papers that summarise other papers (systematic reviews and meta-analyses) *BMJ* 315: 672-675
- Guyatt GH, Sinclair J, Cook DJ, Glasziou P; 1999. Users' Guides to medical literature: XVI. How to use a treatment recommendation. *JAMA* 281(19):1836-1843.
- Rosenfeld RM; 1996. How to systematically review the medical literature. *Otolaryngology-Head and Neck surgery* 115(1):53-63.
- Oxman AD, Cook DJ, Guyatt GH; 1994. Users Guides to the Medical Literature: VI. How to use an overview. *JAMA* 272(17): 1367-1371.
- Sackett DL, Richardson WS, Rosenberg W, Haynes RB; 1997. Evidence-Based Medicine: How to practice and teach EBM. Churchill-Livingstone, First publication.