Lex Bouter and Gerben ter Riet
Department of Epidemiology and Health Care Research, University of Limburg, P.O. Box 616, 6200 MD Maastricht, the Netherlands

Knowledge about the efficacy of the numerous different treatments offered by physiotherapists is essential for patients, health care professionals (including physiotherapists) and policymakers. Elsewhere in this volume (Bouter et al 1990) it is explained why the efficacy of physiotherapeutic interventions should preferably be studied with randomized clinical trials (RCTs). Meta-analysis refers to a critical summary of the evidence on the efficacy of an intervention that is based on the available RCTs on the topic. Meta-analyses usually focus either on the methodological quality of available trials, or on providing a quantitative summary of the effect (pooling). Both forms differ from traditional 'impressionistic' reviews by their explicit and standardized methods. Where traditional review articles tend to rely heavily on the authority of the reviewer, in meta-analyses the reader is, to a substantial extent, able to draw his or her own conclusions, which may differ from those of the author(s).

Meta-analyses seem to have become rather popular recently. The methodology originally was developed by social scientists (see for example Light & Pillemer 1984) in the seventies, and has later been adapted for use in the medical field (Greenland 1987, Sacks et al 1987, Boissel et al 1989). Until now the emphasis has primarily been laid upon the pooling of RCTs, but growing attention is now being given to a standardized methodological assessment of the studies involved (Chalmers et al 1981, Goldschmidt 1986, Anonymous 1987, Gerbarg & Horwitz 1988). For the field of physiotherapy, the latter form, sometimes labelled as 'criterium-based meta-analysis', is probably the most valuable. The reason for this is that for most physiotherapeutic treatments only a relatively small number of RCTs will be available, with a large variation in methodological quality. Consequently, the main topic of this article will be the methodology of criterium-based meta-analyses.

Firstly, the general aims of meta-analyses will be described. Secondly, potential sources of bias in the study of literature are presented. Thirdly, some crucial aspects of the protocol of a meta-analysis are explained. Finally, examples of current and future use of meta-analysis in the field of physiotherapy will be discussed.

Aims

The purpose of a meta-analysis can be both clinical and scientific. The clinical relevance consists of the possibility of getting a more precise estimate of the effect of an intervention. This is especially important when the available trials have a low power, which means that each trial has only a small chance of detecting an existing difference that would be considered to be of clinical relevance. In a meta-analysis the sensitivity for such a small, but clinically relevant, effect will be elevated substantially. This can be formalized in a pooled statistical analysis (Yusuf et al 1985, Demets 1987), but the details of this technique fall outside the scope of this article. Another related advantage of a meta-analysis consists of the improved possibilities (compared to those in individual trials) of getting an impression of the effect of the intervention among different groups of patients, and of the influence of doses, duration of treatment and co-interventions on the efficacy of the treatment at issue.

The scientific relevance of a meta-analysis consists of an explicit summary of the 'state-of-the-art'. In addition to providing an overall judgement with respect to efficacy, the meta-analysis focuses attention on gaps and methodological weaknesses of available studies. Consequently, guidelines can be developed for the design of future RCTs with the aim of answering remaining questions. Meta-analyses are increasingly used for policy decisions, for instance concerning the registration of a new drug (Furberg & Morgan 1987) or the coverage of a new treatment by health care insurances (Louis et al 1985).

Sources of bias

A study of literature is by definition non-experimental and therefore open to the same forms of bias as other observational studies: selection bias, information bias, effect modification and confounding (see for example Rothman 1986).

Not all RCTs are carried out, analyzed and published according to the original planning. Furthermore, 'negative' studies seem to have a substantially smaller chance of appearing in print, especially when they have a small sample size (Dickersin et al 1987). In other words, small trials that are published will have an elevated chance of being 'false positives'. This will lead to selection bias on the level of the meta-analysis by causing an overestimation of the efficacy of the treatment at issue. This phenomenon is called publication bias for which correction is not easy to perform. The only real solution to this problem consists of registration of all trials when they are started, so that later both published and unpublished results can be included in a meta-analysis (Simes 1987).

Although the performance of RCTs forms a necessary condition for the valid assessment of the efficacy of an intervention (Bouter et al 1990), this is by no means a sufficient condition. The (internal) validity of RCTs is often seriously impaired (Chalmers et al 1981). Furthermore, there are indications that trials of doubtful methodological quality generally tend to report positive results more often (Louis et al 1985, Ter Riet et al 1990). Inclusion of all RCTs in a meta-analysis, not taking into account the (internal) validity of the trial, could introduce a substantial amount of information bias leading to an overestimation of the efficacy of the treatment at issue. This constitutes a serious problem, especially in a field where the methodological standards of the RCTs are (still) rather low. An example of this is provided by two recent meta-analyses of the efficacy of acupuncture in the treatment of chronic pain: a pooled analysis of 14 of all available (N=51) trials (Patel et al 1989) appeared to lead to a more optimistic overall judgement than a criterium-based meta-analysis that took into account the methodological quality of the trials involved (Ter Riet et al 1990).

Although there is nowadays more or less consensus about the criteria for an (internally) valid RCT (Bouter et al 1990), the application of these criteria and the weights given to a violation of each of them remains arbitrary to some extent. Therefore, it is important that standardized and explicit criteria are used by several expert reviewers who assess the trials independently and are blinded for the results of the trial, its authors and the journal in which it appeared (Chalmers et al 1987, Furberg and Morgan 1987, Gerbarg and Horwitz 1988). The reviewers should have a high level of agreement and are to be recruited both from clinical experts (physiotherapists) and experts in research methodology (epidemiologists or biostatisticians).

Another important issue in the design of a meta-analysis consists of the comparability of patients and interventions among the RCTs involved. The question here is whether the existing differences between the groups of patients and the execution of the interventions (including co-interventions) will have an influence on the efficacy of the treatment at issue. Presence of such an influence will lead to effect modification on the level of the meta-analysis. The solution to this depends on which question the meta-analysis tries to answer. Often restriction of the meta-analysis to a homogeneous subgroup of patients or to a specific operationalization of the intervention will be indicated. But when the purpose is to study the influence of a prognostic factor on the efficacy of the intervention, the meta-analysis should be divided in clusters corresponding with this factor (e.g. by presenting the efficacy of ultra sound therapy separately for acute and for chronic low back pain patients).

Incomparability of effect parameters among the trials involved constitutes another source of bias which can confound the overall judgement of efficacy (Greenland 1987). The outcome parameter for which the RCTs are compared should be identical (or at least very similar) in all trials involved. Often some form of standardization, for example the percentage of pain reduction, will be necessary to enable comparison.

Protocol for a meta-analysis

Like every empirical study, a meta-analysis needs an explicit research protocol that is formulated before the review of the literature is executed. This protocol will state the central question the meta-analysis tries to answer, and contains details about inclusion and exclusion criteria for trials and the items and weights of the methodological assessment. Furthermore, the protocol will provide information on the way standardization and blinding of the procedure will be guaranteed. Because of differences in subjects and aims, every meta-analysis will need its own protocol (Louis et al 1985, Sacks et al 1987, Furberg and Morgan 1987). Therefore, this section can only present some general guidelines.

The inclusion and exclusion criteria deal primarily with the patients, interventions and effect parameters that are eligible. In addition to this, they can contain, for instance, a minimal duration of follow-up, or mention a period in which the study has to be executed or published. Next, the method of identification of eligible trials has to be stated explicitly. Usually computerized data bases (e.g. Index Medicus, Excerpta Medica and Current Contents) are used as a starting point. This will by no means identify all eligible trials, because the key words of these systems are not very appropriate for physiotherapeutic purposes and not all journals relevant to physiotherapists are abstracted. Therefore, additional sources should be used, like textbooks, earlier reviews, lists of references and\correspondence with experts in the field (Sacks et al 1987). Especially if one wants to minimize publication bias by including unpublished data as well, sending a questionnaire to all authors of relevant publications already identified can be very helpful.

An important part of the protocol deals with the methodological assessment of the trials. The reviewers of the methodological quality of the trials ought to be selected and they will have to agree upon the items they will focus on. These items will partly deal with methodological issues that are central in all RCTs (e.g. method of randomization, loss to follow-up or use of a placebo treatment; see for instance Fletcher et al 1988). Next to these, it is usual to add a number of items (e.g. dealing with the choice of the effect parameters, the execution of the interventions or relevant variables for the assessment of prognostic comparability) which are more specific for the subject at issue. When a certain weight is allocated to the items, for each trial an overall score for methodological quality can be calculated (Chalmers et al 1981, Ter Riet et al 1990).

The protocol of the meta-analysis should state clearly whether and for what aspects (results, authors or journal) the methodological assessment will be blinded. Of course, blinding has to be realized by a person not involved in reviewing the trials. Unfortunately, reading the methods section only will usually not be sufficient, because relevant methodological details may be found in other sections of the article as well. After an initial assessment of the trials by the reviewers, it may be necessary to operationalize some items further before the overall methodological score can be calculated. The publication in which the results of the meta-analysis are presented should preferably specify the level of agreement between the reviewers and enable the reader to generate his or her own judgement by applying other weights to the methodological items.

Discussion

The principles of reviewing the literature on the efficacy of a physiotherapeutic treatment that are mentioned in this article are relatively new and by no means fully developed. Criterium-based meta-analyses as advocated above are still not very common (Dickersin et al 1987, Gerbarg & Horwitz 1988). Most meta-analyses published in the medical field focus on generating a quantitative summary (pooling) of all available evidence, largely ignoring differences in methodological quality of the trials at issue (Sacks et al 1987). We think that for the situation in the physiotherapeutic literature the latter would be an example of bad practice. The reason for this is that RCTs on the efficacy of a physiotherapeutic treatment tend to vary substantially with respect to methodological quality. Moreover, for a given question of efficacy in the physiotherapeutic realm usually only a limited number of RCTs (0-20) are available, so there is no need for pooling the results.

Recently, a number of criterium-based meta-analyses were performed at our department. One of them deals with the efficacy of acupuncture in the treatment of chronic pain and is presented elsewhere in this volume (Ter Riet et al 1990). Currently we are executing a large literature study on the efficacy of physiotherapeutic interventions, including treatment modalities like electro therapy, ultra sound, traction, laser therapy and manipulative therapy, all for several indications. In this project the idea of criterium-based meta-analyses will be tested critically and developed further with special attention to feasibility and informativity.

References

Anonymous (1987), Whither meta-analysis? [Editorial] Lancet, i, 897-898

Boissel JP, Blanchard J, Panak E, Peyrieux JC and Sacks H (1989), Considerations for the meta-analysis of randomized clinical trials. Controlled Clinical Trials, 10, 154-281

Bouter LM, Linden SJ van der and Koes BW (1990), How to assess the effects of physiotherapy: a critical review of the methodology with special attention to the outcome parameters. In this volume

Chalmers TC, Levin H, Sacks HS, Reitman D, Berrier J and Nagalingam R (1987), Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. Statistics in Medicine, 6, 315-325

Chalmers TC, Smith H Jr, Blackburn B, Silverman B, Schroeder B, Reitman D and Ambroz A (1981), A method for assessing the quality of a randomized control trial. Controlled Clinical Trials, 2, 31-49

Demets DL (1987), Methods for combining randomized clinical trials: strengths and limitations. Statistics in Medicine, 6, 341-348

Dickersin K, Chan S, Chalmers TC, Sacks HS and Smith H Jr (1987), Publication bias and clinical trials. Controlled Clinical Trials, 8, 343-353

Fletcher RH, Fletcher SW and Wagner EH (1988), Clinical epidemiology: the essentials. Williams & Wilkins, Baltimore

Furberg CD and Morgan TM (1987), Lessons from overviews of cardiovascular trials. Statistics in Medicine, 6, 295-303

Gerbarg ZB and Horwitz RI (1988), Resolving conflicting clinical trials: guidelines for meta-analysis. Journal of Clinical Epidemiology, 41, 503-509

Goldschmidt PG (1986), Information synthesis: a practical guide. Health Services Research, 21, 215-237

Greenland S (1987), Quantitative methods in the review of epidemiologic literature. Epidemiologic Reviews, 9, 1-30

Light RJ and Pillemer PB (1984), Summing up: the science of reviewing research. Harvard University Press, Cambridge

Louis TA, Fineberg HV and Mosteller F (1985), Findings for public health from meta-analysis. Annual Reviews of Public Health, 6, 1-20

Patel M, Gutzwiller F, Paccaud F and Marazzi A (1989), A meta-analysis of acupuncture for chronic pain. International Journal of Epidemiology, 18, 900-906

Riet G ter, Kleijnen J and Knipschild PG (1990), Acupuncture and chronic pain: a criteria-based meta-analysis. In this volume

Rothman KJ (1986), Modern Epidemiology. Little, Brown & Co, Boston

Sacks HS, Berrier J, Reitman UA, Berk A and Chalmers TC (1987), Meta-analysis of randomized controlled trials. New England Journal of Medicine, 316, 450-455

Simes RJ (1987), Confronting publication bias: a cohort design for meta-analysis. Statistics in Medicine, 6, 11-29

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