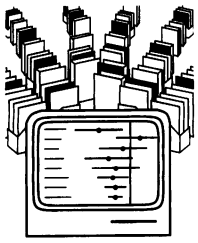
Obtaining data from randomised controlled trials: how much do we need for reliable and informative meta-analyses?

Michael J Clarke, Lesley A Stewart



This paper was presented at a meeting of Systematic Reviews organised jointly by the BMJ and the UK Cochrane Centre and held in London in July 1993

Many randomised controlled trials compare treatments that will produce only moderate differences in outcome, but these differences can be clinically important. However, they are difficult to assess reliably and require a large amount of randomised evidence. This can be achieved through large prospective randomised trials which will accrue future patients, the meta-analysis of results from randomised trials involving patients from the past, or—ideally—both. The techniques require that all possible biases are minimised, and in meta-analyses this can best be achieved by ensuring that all of the randomised evidence—both trials and participants in those trials—is included. The meta-analysis of individual patient data has been described as the gold standard for this approach. It will remove many of the problems associated with relying solely on published data and some of the problems arising from a reliance on aggregate data, and will also add to the analyses that can be performed. Such projects, however, require considerable time and effort.

The differences in outcome between many of the treatments compared in randomised trials are moderate but potentially very important to patients and their medical carers. Individually, however, most trials have been too small to assess such differences reliably. There are two main ways to overcome this: through large prospective randomised trials which will accrue future patients, and through meta-analyses of completed trials. Whether a single randomised trial or a meta-analysis is to be undertaken, all possible biases should be minimised, and perhaps the most important step in this is to ensure that as much as possible of the randomised evidence is included in the analyses.[1]

This paper sets out the reasons for and the means of doing this. It emphasises meta-analysis using individual patient data, which has been described as a yardstick against which other forms of systematic review could be measured,[2] but many of the points raised are also relevant to meta-analyses using aggregate data. The individual patient data approach requires that data on every patient entered to all relevant randomised trials are collected centrally, allowing careful data checking and standard analyses to be performed and an overall result, based on the totality of the available evidence, to be calculated. These projects can provide reliable evidence in areas of uncertainty, as in the Early Breast Cancer Trialists' Collaborative Group's meta-analyses of randomised trials of tamoxifen and chemotherapy[3,4] and the Non-Small Cell Lung Cancer Collaborators Group's meta-analyses of chemotherapy.[5]

Several other such projects have been undertaken successfully and the little quantitative and empirical evidence published to date has shown their advantages over reviews based on published or aggregate data alone.[6,7] These advantages arise because of the increased accuracy and updating of the material available for review and the additional analyses that are possible with individual patient data but cannot be done with aggregate data alone. Some of the advantages of collecting individual patient data are described below, along with suggestions on how such data may be collected. This work is based on the practical experience of two of the groups who have acted as secretariats for some of the largest meta-analyses based on individual patient data conducted to date. (To expand this experience, a workshop was organised recently by the Cochrane Collaboration to bring together representatives of other groups undertaking such projects. A full report of the findings of the workshop, including areas such as protocol use and development, methods of checking data, and resource requirements is being prepared.)

Minimising biases and random errors

COMPLETE IDENTIFICATION OF PUBLISHED AND UNPUBLISHED TRIALS

The most important step in the conduct of any systematic review of randomised controlled trials is to identify and include all (or nearly all) of the relevant trials. This is needed whether the review is to be based on aggregate or individual patient data; the process of trial identification has already been described in this series.[8] Meta-analyses based on individual patient data always require direct contact with trialists (as do some reviews based on aggregate data), so these provide an additional means of identifying trials—enlisting the help and knowledge of those trialists. For example, neither of two important reviews of randomised trials comparing melphalan and prednisone with combination chemotherapy in the treatment of multiple myeloma[9,10] found the unpublished Italian M-80 randomised trial of these drugs versus vincristine plus melphalan, cyclophosphamide, and prednisone. This study was also unknown to the secretariat of an ongoing overview of such trials until the Italian group was contacted for patient data from its other trials. Similarly, direct contact with trialists identified two unpublished trials previously unknown to the secretariat of a meta-analysis of advanced ovarian cancer[11]; these had not been identified by a meta-analyses based on the published literature.

OBTAINING INFORMATION ON ALL RANDOMISED PARTICIPANTS AND EXCLUDING INFORMATION ON THOSE WHO WERE NOT RANDOMISED

All randomised patients should be included in the analysis in accordance with the treatment allocated at randomisation (an "intention to treat" analysis). In this way, the policy of using one treatment will be appropriately compared with the policy of using another.[12] Unfortunately, many randomised trials do not follow this principle when publishing their results, and patients are excluded for a variety of reasons. Sometimes these reasons will seem unconnected with the assigned treatment—for example, when the delayed result of a prerandomisation diagnostic test reveals that the patient was ineligible for the study. More seriously, the reasons can be related to treatment—for example, the patient may have been unable to tolerate the allocated treatment or failed to follow the treatment schedule for some other reason.

Many published papers will state that some patients have been judged ineligible and omitted from the

Clinical Trial Service Unit and ICRF Cancer Studies Unit, Radcliffe Infirmary, Oxford OX2 6HE
Michael J Clarke, *cancer overviews coordinator*

MRC Cancer Trials Office, Cambridge CB2 2BW
Lesley A Stewart, *overview coordinator*

Correspondence to:
Dr Clarke.

BMJ 1994;309:1007-10

analyses, and the people reviewing the paper will be aware of the size of the problems this might cause. A more difficult problem arises if a publication contains no mention of randomised but ineligible patients, usually because the trialist considers that these patients are no longer part of the trial. Such a situation was noted recently by Hoover *et al* for their randomised trial of active-specific immunotherapy in colorectal cancer. The second publication[13] of this study's results noted that an oversight in an earlier publication[14] led to the failure to state that some patients were randomised but excluded from the analyses.

Occasionally, some non-randomised patients are included in a trial's published analyses—then it is important that these patients are excluded from the meta-analysis. This can happen if a randomised trial was preceded by a non-randomised run in phase, or if patients continue to be entered to one of the study's treatments after the randomisation has been closed. It can also happen if the randomisation is temporarily stopped during the trial. Figure 1 shows this for an unpublished trial of radiotherapy versus chemotherapy in multiple myeloma. The radiotherapy equipment was not available for six months during the trial, but patients continued to enter the chemotherapy arm. The appropriate analysis of the trial would exclude this group of patients, but it was only when the data were supplied for the overview that the problem was brought back to the attention of the trialist.

OBTAINING COMPLETE AND UNBIASED INFORMATION ON ALL SUBGROUPS AND OUTCOMES STUDIED

A trial that collects information on a variety of patient characteristics can have as many subgroup analyses as there are types of patient in the data. Whether or not these analyses provide useful information will not be debated here. Constraints on space and other influences make it most likely that the analyses relating to the subgroups with the most striking results will be published. Thus, any subgroup analyses in a systematic review that uses only those subgroups available in published reports will be subject to both the effect of publication bias in the trials available for inclusion and an additional bias in the subgroups available for analysis.

Similarly, if a trial measures several outcome measures there will be a tendency for those showing the most striking results to be published. This could occur if a series of alternative measures is used for the same outcome, such as rating scales in psychiatric illness, or if the outcome measure can vary depending on the convention used to define it, such as event free survival in leukaemia.

The collection of unpublished material helps with this in two main ways. If the patient characteristics are obtained for all trials, subgroup analyses based on the total evidence can be performed. Moreover, these analyses will contain a larger number of events and have greater statistical power than any single trial.

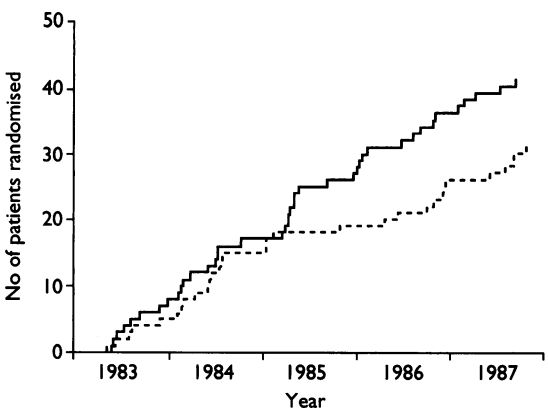


FIG 1—*Entry of patients to randomised trial showing accrual of patients to chemotherapy (and radiotherapy) treatment groups. Individual patient data include patients entered to the chemotherapy group during April to September 1985, when radiotherapy was not available, these patients should not be included in the meta-analysis. (Figure included with permission of trialist)*

With regard to outcome measures, it may be possible to specify a uniform definition for a particular outcome and analyse this across all trials. It is worth remembering that whether or not a subgroup or outcome can be analysed will depend on its initial collection by the trialist (a decision that could not have been biased by the trial's own results) and on the willingness of the trialist to supply data on that variable—a problem that can happen with any of the data or trials in the meta-analysis.

OBTAINING COMPLETE FOLLOW UP DATA

Whenever the results of a trial are published, those results become "frozen in time" and will usually remain so unless the trial is updated and published again, which happens rarely. The supply of information for a meta-analysis is another way in which the trial's results can be updated, both by providing additional follow up and by completing data that were missing at the time of publication. The effect of the additional material on the results of a meta-analysis will vary. For example, an overview of the comparison of single non-platinum drugs with platinum-based combination chemotherapy in advanced ovarian cancer found that increasing the period of follow up reduced the estimate of the overall treatment benefit.[6] In contrast, the breast cancer overview found an additional benefit for patients allocated to adjuvant chemotherapy, rather than control, when the follow up was extended and standardised to the period 5-10 years after treatment.[4] This result was so surprising to the statistical secretariat at the time of the preliminary analyses that a questionnaire was sent to all of the collaborating trialists. Seventy eight replied and their predictions for the additional effect of prolonged multiagent chemotherapy in premenopausal women during years 5-10 after surgery ranged from an increase in the odds of death of 20% to a decrease of 25%. None were as extreme as the overview result—additional decrease of 33%.

HOW TO OBTAIN DATA THAT ARE AS COMPLETE AS POSSIBLE

Whether the information on the participants in the relevant randomised trials should be collected as aggregate data or individual patient data will be discussed briefly below. In either case, it must be collected from as many trials as possible. It is especially important to ensure that any trials that do not contribute data are not so numerous or unrepresentative to introduce important bias into the result of the systematic review. Thus the data collection process may present the reviewer with several difficulties. Some trialists may be reluctant to supply their data, and there will often be practical difficulties in preparing data. It is important therefore to emphasise that any data supplied will be treated confidentially and will not be used for any additional purpose without the permission of the responsible trialist. In addition, any publications arising from the meta-analysis should be in the name of all the collaborators, and each trialist should have an opportunity to comment on the manuscript before publication. The trialists will also be the first people, other than the statistical secretariat, to see and discuss the overview results if these are presented first to a closed meeting of the collaborative group of all participating trialists.

If there are trialists who initially were unable to prepare and supply their data, some of these points may help persuade them to do so. In addition, the process of data collection should be as simple and flexible as possible so as to help and encourage trialists to participate. In some instances, even if the initial request was for aggregate data it may be easier and preferable for the trialist to supply individual patient

data so that the necessary tables can be prepared centrally. This might be the case if, for example, only paper records exist for each patient in the trial or if the trialist does not have the necessary resources to prepare the tables.

Patience also helps because data that are regarded as completely lost may sometimes reappear. For example, the third cycle of the early breast cancer overview will include a trial whose records were feared lost in a flood at the time of the second cycle, but which were recently found when an office was cleared.

Benefits of using individual patient data rather than aggregate data

The reasons for obtaining information on all randomised patients in all relevant trials have been described above, and these goals can often be achieved by collecting aggregate information from the responsible trialists. Collecting individual patient data may allow some of them to be achieved more easily or reliably and will also provide important additional benefits.

CALCULATION OF TIMES TO EVENTS

Perhaps the most substantial benefit is that it is not possible to calculate and analyse the times to specific events reliably without individual patient data. Such analyses might reveal prolongation of event free periods or differences in median survival for the treatments being compared. Figure 2 shows hypothetical examples of how the divergence of the survival curves, which might be of great relevance to patients and clinicians, would be missed if aggregate data were collected for any time beyond point B. Data collected only for time A would produce an overoptimistic conclusion on the treatment benefit in the upper example. In addition, the time to event analyses contribute greater statistical power than is possible with the limited number of time points that would be available with aggregate data.

CHECKING AND CORRECTING DATA

Although our experience after looking at data from hundreds of trials is that deliberate fraud in randomised controlled trials is rare, the requirement for individual patient data can help to serve as a check on the use of fabricated data, either from a complete trial or part of a trial. Much more commonly, the central review of patient data will highlight problems with a randomised trial that occurred through error rather than fraud, these mostly arise during the process of randomisation itself or in the follow up of patients in the treatment groups.

For example, the individual patient data might reveal the exclusion of randomised participants or the inclusion of some who were not randomised. This would come to light if the data provided contain a different number of patients from that reported for a trial or if there were sequence gaps due to the absence of some patients. In either case the data on the missing patients could then be requested from the trialist, or the inclusion of additional non-randomised patients could be queried. The patient data will also reveal if patients who were inappropriately classified as ineligible have as much follow up information as eligible patients. If not, the trialist can be asked for further information. Once all of the data are available, the appropriate intention to treat analysis can be performed.

SUPPLY OF ADDITIONAL PATIENT DATA

If individual patient data are collected it is relatively easy for a trialist to supply additional follow up information or previously missing data on selected

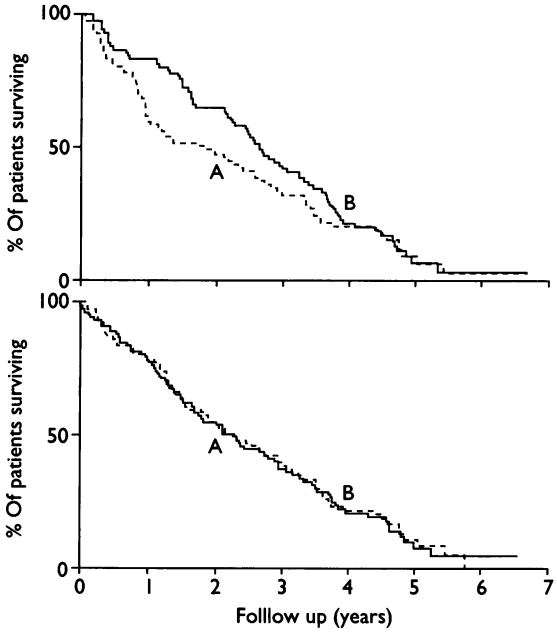


FIG 2—*Simulated survival curves with approximately the same survival in both groups at 4 years (B) but not at 2 years (A)*

patients and for this to be incorporated in the meta-analysis, but if aggregate data were collected the trialist would have to produce a new set of tables. In meta-analyses where death is a major end point, the individual patient data may also allow the central secretariat to continue the follow up through national death registers.

MORE FLEXIBLE AND POWERFUL ANALYSIS OF SUBGROUPS AND OUTCOME MEASURES

If subgroup analyses are to be performed and different outcome measures are to be used, the complexity of the summary tables might become such that they resembled individual patient data. For example, having seven important categories of patient, two important outcome measures and a randomisation into two groups would result in a table with 28 separate values. This is the minimum that would have been required for one table of results in the report of the overview of systemic treatment versus control in breast cancer, which contained the recurrence free and overall survival results among women above and below 50 years of age, with and without nodal involvement, who had positive, negative, or unknown oestrogen receptor status.[4] In addition, because the outcome data will almost certainly be needed for more than one point in time, the aggregate data must be produced for each of these points.

OTHER ADVANTAGES OF THE COLLABORATIVE EFFORT

The involvement of a group of trialists in a meta-analysis can provide wide experience and helpful input when the results are being prepared for publication. The effort involved in collecting and analysing the data can justify holding a collaborators' meeting at which this experience can be expressed and assimilated in a much more interactive way than is possible with the circulation of a draft manuscript. It also allows an additional check that each trialist's data have been properly included in the meta-analysis.

Conclusion

The most important first step in any systematic review is ensuring that all, or nearly all, relevant trials are identified. After that, the data for analyses can be gathered in a variety of ways. Collecting individual patient data centrally is perhaps the most resource intensive and time consuming of these. This will,

however, overcome many of the problems associated with relying solely on published data and some of the problems associated with relying on aggregate data and will add to the analyses that can be performed. It might therefore provide the "gold standard" to which systematic reviews should strive.[15]

Which steps in the process are the most important for improving reliability requires further testing and evidence, especially if some of these steps lengthen the time needed to conduct the meta-analysis but do not greatly improve its reliability. To this end, some of the topics for consideration would be the use of trials from which individual patient data are not available but published data are and of trials in which the individual patient data reveal problems (such as the inappropriate exclusion of some patients and the subsequent destruction of their relevant records) that cannot be rectified.

Just as different forms of health care need to be reliably assessed, so the techniques for reviewing evidence from randomised controlled trials should be empirically investigated.

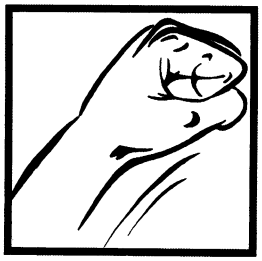
1 Peto R, Collins R, Gray R. Large scale randomized evidence: large, simple trials and overviews of trials. *Ann N Y Acad Sci* 1993;**703**:314-40.
2 Chalmers I. The Cochrane Collaboration: preparing, maintaining, and disseminating systematic reviews of the effects of health care. *Ann N Y Acad Sci* 1933;**703**:156-65.
3 Review of mortality results in randomised trials in early breast cancer. *Lancet* 1984;ii:1205.
4 Early Breast Cancer Trialists' Collaborative Group. Systematic treatment of early breast cancer by hormonal, cytotoxic or immune therapy: 133 randomised trials involving 31,000 recurrences and 24,000 deaths among 75,000 women. *Lancet* 1992;**339**:1-15. 71-85.
5 Stewart LA, Pignon JP, Arriagada R, Souhami RI, on behalf of the NSCLC Collaborative Group. A meta-analysis using individual patient data from randomised clinical trials of chemotherapy in non-small cell lung cancer. I. Survival in the surgical setting [abstract]. *Proceedings of the American Society of Clinical Oncology* 1994;**13**:336.
6 Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;**341**:418-22.
7 Pignon JP, Arriagada R. Meta-analysis. *Lancet* 1993;**341**:964-5.
8 Dickersin K, Scherer R, Lefebvre C. Identification of relevant studies for systematic reviews. *BMJ* (in press).
9 Simes J. Publication bias: the case for an international registry of clinical trials. *J Clin Oncol* 1986;**4**:1529-41.
10 Gregory WM, Richards MA, Malpas JS. Combination chemotherapy versus melphalan and prednisolone in the treatment of multiple myeloma: an overview of published trials. *J Clin Oncol* 1992;**10**:334-42.
11 Advanced Ovarian Cancer Trialists Group. Chemotherapy in advanced ovarian cancer: an overview of randomised clinical trials. *BMJ* 1991;**303**: 884-9.
12 Peto R, Pike M, Armitage P, Breslow NE, Cox DR, Howard SV, *et al.* Design and analysis of randomised clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976;**34**:585-612.
13 Hoover HC, Brandhorst JS, Peters LC, Surdyke MG, Takeshita Y, Madariaga J, *et al.* Adjuvant active specific immunotherapy for human colorectal cancer: 6·5-year median follow-up of a phase II prospectively randomised trial. *J Clin Oncol* 1993;**11**:390-9.
14 Hoover HC, Surdyke MG, Dangel RB, Peters LC, Hanna MG. Prospectively randomised trial of adjuvant active-specific immunotherapy for human colerectal cancer. *Cancer* 1985;**55**:1236-43.
15 What's wrong with the BMJ [editor's choice]. *BMJ* 1994 January 8;**308**:i.

Controversies in Management

Are antibiotics appropriate for sore throats?

Costs outweigh the benefits

P S Little, I Williamson



This is the seventh in a series of articles examining some of the difficult decisions that arise in medicine

Department of Primary Medical Care, Aldermoor Health Centre, Southampton University, Southampton SO1 6ST
P S Little, *Wellcome training fellow*
I Williamson, *senior lecturer*

BMJ 1994;**309**:1010-2

General practitioners prescribe antibiotics for sore throat for various reasons including to prevent complications (rheumatic fever, glomerulonephritis, sinusitis, otitis media, etc), to relieve symptoms, and for psychosocial reasons. However, the benefit is marginal and the costs are great.

Do antibiotics prevent complications?

Studies on the prevention of rheumatic fever were carried out using penicillin injections in military personnel in barracks after the second world war.[1] The attack rates were high (0·3-5%), and the results may not be generalisable to a modern community setting with lower attack rates and where the likelihood of developing rheumatic fever or glomerulonephritis is the same in those who have and have not had oral antibiotics.[2,3] The incidence of rheumatic fever has been falling since the turn of the century—well before antibiotics were discovered.[4] General practitioners in Britain have about a one in five chance of ever seeing a patient with either post-streptococcal glomerulonephritis or rheumatic fever after a sore throat.[2,3]

The main problem of prescribing to prevent these problems is that most patients with sore throat never attend their general practitioner.[2,3,5] Even if the benefit of oral antibiotics in the community were proved general practitioners' surgeries would need to be overwhelmed with patients or antibiotics would need to be freely available in the community to prevent such complications effectively.

Some evidence exists for a small protective effect of antibiotics on the development of otitis media and sinusitis.[1] However, these studies are old, included small numbers of complications, and were mainly conducted in institutionalised servicemen. Studies in general practice had very wide confidence intervals for the odds ratio for developing complications (greatly overlapping 1 for prevention of otitis media).[6,7] Thus it seems doubtful whether oral antibiotics prevent suppurative complications of sore throat. Even if large modern studies supported these results at least 29 subjects with sore throat would have to be treated to prevent one case of otitis media,[1] which is usually a self limiting condition.

Other reasons for prescribing

The evidence for relief of symptoms in sore throat is also marginal. Results from the few placebo controlled trials in general practice suggest there may be a small increase in the number of patients well after three days among those taking penicillin.[1] However, the largest trial (n=528) showed this benefit for only a small subgroup of the study population.[8] Furthermore, the illness was not shortened at all irrespective of initial presentation with fever, purulent tonsils, or lymphadenitis (figure).

Psychosocial factors for both the doctor and the patient are important determinants of prescribing,[9,10] and it is important to acknowlege and explore them. General practitioners probably perceive more pressure to prescribe than exists, since 41% of patients entering consultations expect a prescription but 67% leave with one.[11] Even if patients receive an antibiotic for sore throat a 10 day course would be needed to eradicate streptococci, and the evidence suggests that only half of children complete such a course.[12] An uncontrolled