

Preparing and Updating Systematic Reviews of Randomized Controlled Trials of Health Care

IAIN CHALMERS, MURRAY ENKIN,
and MARC J.N.C. KEIRSE

*The UK Cochrane Centre, United Kingdom;
McMaster University, Canada;
University of Leiden, The Netherlands*

If, as is sometimes supposed, science consisted in nothing but the laborious accumulation of facts, it would soon come to a standstill, crushed, as it were, under its own weight. The suggestion of a new idea, or the detection of a law, supersedes much that has previously been a burden on the memory, and by introducing order and coherence facilitates the retention of the remainder in an available form. . . . Two processes are thus at work side by side, the reception of new material and the digestion and assimilation of the old; and as both are essential we may spare ourselves the discussion of their relative importance.

From an address to the British Association given by Lord Rayleigh in 1884

THE IMPORTANCE OF SYNTHESIZING NEW AND existing research evidence has been long recognized. In spite of Lord Rayleigh's injunction, most scientists operate on a double standard: they go to great lengths to define the methods they used to minimize biases and random errors in their reports on the results of new research, but they often do not attempt to apply scientific principles in their discussions of how the newly generated evidence accords with previously available information. Scientists also operate by this double stan-

dard when they conduct and report “stand alone” reviews: an analysis of review articles published in four major medical journals concluded that “current medical reviews do not routinely use scientific methods to identify, assess, and synthesize information” (Mulrow 1987).

This shortcoming is serious. People providing health care rely more on reports of reviews (secondary research) than on reports of primary research to learn about the effects of care (Williamson et al. 1989). Thus, failure by reviewers to apply scientific principles to this secondary research can have adverse consequences for patients, and can lead to waste of health service resources.

In our field, for example, influential reviewers have repeatedly made false positive and false negative inferences about the effects of administering corticosteroid drugs to women expected to deliver preterm because they failed to recognize, acknowledge, and control random errors (the play of chance). As a result, tens of thousands of preterm babies have been denied a highly effective form of care, thousands have suffered and died unnecessarily (Crowley et al. 1990), and the costs of neonatal care have been unnecessarily high (Mugford, Piercy, and Chalmers 1991).

Failure to use corticosteroids is not the only example of the adverse effects of scientifically invalid reviews. A recent study of treatment recommendations made in medical textbooks showed that advice on some life-saving therapies was delayed for more than ten years because scientific principles were not applied to the review process. Other treatments continued to be recommended long after controlled research had demonstrated them to be either ineffective or actually harmful (Antman et al. 1992).

If discoveries about the effects of health care are to be reflected in improved health and more efficient use of resources, the key role of reviews in the pathway between primary research and improved health outcomes must be more widely appreciated. Methodologically sound reviews are essential to guide the providers and consumers of health care, and to identify priorities for the health care research agenda.

Control of Biases in Reviewing the Effects of Care

Judgments about the value of particular forms of care must draw on different forms of evidence (Chalmers 1989). Sometimes the consequences

of policies are obvious, and are unlikely to be confused with the effects of bias: examples include the beneficial effects of ventricular defibrillation after cardiac arrest and the teratogenic effects of thalidomide. Much more commonly, policies and practices have less dramatic, but still important, beneficial or harmful effects, which can only be validly assessed with evidence from studies designed to minimize the effects of bias. This usually means evidence from randomized controlled trials (RCTs).

The increasing recognition of the need for methodological stringency during the past decade unfortunately has only just begun to be reflected in the way research is reported in journals and in the coding practices used by bibliographic databases like MEDLINE. It is therefore still difficult for reviewers to identify the studies that are likely to be methodologically sound, and thus eligible for inclusion in their reviews (Dickersin et al. 1985; Chalmers et al. 1989b).

Bias in secondary research (reviews) can be avoided only by considering all the relevant evidence. Sometimes a sufficiently large, well-controlled trial can provide enough information to allow an informed policy decision. The results of a recently reported multicenter randomized trial, for example, showed that supplementation with folic acid around the time of conception reduces the risk of a mother giving birth to a second baby with a neural tube defect (anencephaly or spina bifida) (Medical Research Council Vitamin Study 1991). The results of this one trial were sufficiently convincing to prompt the Department of Health in England to issue policy advice to health service workers (Acheson and Poole 1991). Similarly, a single controlled trial demonstrated that treatment with a drug as simple and inexpensive as low-dose aspirin can reduce by a fifth the likelihood of premature death after admission to hospital for heart attack (ISIS-2 Collaborative Group 1988); these results also provide an adequate basis to recommend a policy. Sometimes the results of a single trial may provide sufficient grounds to recommend that a form of care should be abandoned. A newly introduced form of suture material to repair perineal trauma after childbirth was found to double the proportion of women experiencing pain for up to three years after delivery (Grant et al. 1989). This finding made it clear that the new material should not be used.

It is rare, however, for the results of a single study to provide a firm basis for policy. More usually, the effects of health care must be assessed by reviewing the body of evidence generated by a number of controlled trials. It is as important to take steps to control bias during this process

as it is during performance of the primary studies (Chalmers et al. 1989b). This means that the criteria for including studies in the review must be made explicit, that as high a proportion as possible of the studies meeting the criteria (whether published or unpublished) should be identified, and that steps should be taken to minimize biases while assembling data from the eligible studies identified.

For some forms of care, evidence from randomized trials is simply not available. This poses great difficulties in establishing an informed basis for practice or resource distribution. However, care policies and practices that have not been assessed in controlled trials can only be identified by exclusion, after a careful search to find out if controlled trials *do* exist. Such forms of care will merit priority attention in agendas for new research.

Control of Random Errors in Reviewing the Effects of Care

After taking steps to control biases during the process of review, reviewers must try to minimize the risk of being misled by the play of chance. It will always be difficult, and often impossible, to assimilate and synthesize informally numerical data generated by a body of related research studies (Collins et al. 1987). The use of appropriate statistical methods to integrate the results of distinct but similar studies will minimize the risks that reviewers and their readers will be misled by random errors.

Peto (1987) has described the rationale for synthesizing the results of similar but separate randomized trials in "meta-analyses" (also referred to as "overviews," or "pooled analyses"):

While we cannot assume that different trials are exactly comparable, or that patients in different trials are exactly comparable, it is reasonable to assume that if different trials address related questions then there is going to be some tendency for the answers to come out in the same direction. That tendency may well be obscured in individual trials, or even in some cases reversed, by the play of chance. But elsewhere it may remain, and it is that tendency which the overview is trying to detect.

Published reports of meta-analyses became more frequent in the early 1980s. Before 1982, a MEDLINE search could be expected to yield an average of about one meta-analysis a year. Between 1982 and 1985 the average annual yield was about 15 (Dickersin, Higgins, and Meinert 1990). Between 1986 and 1989, the number of meta-analyses listed by MEDLINE more than doubled every year. MEDLINE searches using the MeSH term "META-ANALYSIS" and those using the text word "meta-analysis" yielded, all together, 270 citations for 1990 (K. Dickersin 1991: personal communication).

Many of these meta-analyses have revealed that policy makers and clinicians have incorrectly concluded that certain forms of care are either useless or only useful for certain categories of patients. The predictions of 78 world authorities about the long-term effects of adjuvant therapy in the treatment of breast cancer, for example, were shown in meta-analyses to have seriously underestimated the beneficial effects of tamoxifen and polychemotherapy (Doll 1991).

By contrast, properly conducted meta-analyses can help to identify forms of care, currently being offered and used within the health services, that are either very unlikely to have important beneficial effects or that may actually be harmful. Routine hospitalization of women with twin pregnancies (Crowther 1991) and antiarrhythmic drugs given during myocardial infarction (MacMahon et al. 1988) are examples. Therefore, where appropriate and possible, systematic reviews of the effects of health care should use formal statistical synthesis of the results of separate but similar experiments to control random errors.

Effective Care in Pregnancy and Childbirth: One Example of a Systematic Review of the Effects of Care

For over a decade, assisted by hundreds of people, we have been attempting to improve the quality of reviews of evidence about the effects of care during pregnancy and childbirth. The methods that we used to identify relevant evidence and to commission and conduct reviews of the evidence identified have been reported previously (Chalmers et al. 1986; Chalmers, Enkin, and Keirse 1989a; Hetherington et al. 1989; *International Register of Perinatal Trials* 1991). Here, we will summarize our

methods and comment on the process of commissioning and creating the reviews. In addition, we provide an account of current arrangements for updating and amending reviews as new data have become available and errors have been identified. We conclude with some general reflections on our experiences.

Identifying Relevant Evidence

Appendix 1 provides a chronological summary of the evolution of our work, which originated in 1973 when one of us (IC) read a book of seminal importance that had been published the previous year. Archie Cochrane's *Effectiveness and Efficiency: Random Reflections on Health Services* (1972) was the first really clear exposition of the importance for health services of evidence derived from RCTs. With disarming simplicity and directness, Cochrane asked why the public should be expected to pay (directly or indirectly) for forms of care that had not been shown to be effective either in preventing illness or disability or in improving the natural history of disease.

Under Archie Cochrane's continuing influence, and with his encouragement, we laid plans for developing a register of controlled trials in perinatal medicine. By 1976, we had outlined a plan for systematically reviewing the results of these trials. First, as high a proportion as possible of properly controlled trials would be identified; then the results of similar trials would be synthesized in meta-analyses; and finally, arrangements would be made for updating these analyses continuously as new data became available.

In 1978, with the help of a small but important grant provided by the Maternal and Child Health Unit of the World Health Organization (WHO), we began a systematic hand search of 60 "core" journals initially, starting with the volumes published in 1950. These hand searches supplemented MEDLINE searches of the literature published since 1966 (Chalmers et al. 1986, 1989b). The importance of the hand searches was demonstrated in subsequent analyses, which showed that searches relying on MEDLINE alone resulted in substantial (and possibly biased) underascertainment of relevant studies (Dickersin et al. 1985; Chalmers et al. 1989b).

A year's sabbatical leave in 1979 (ME) enabled two of us (IC and ME) to collaborate in overseeing the initial search for relevant reports, in developing a classification scheme for categorizing them, and, as trials were

identified and classified, in guiding the development of software to enter details about them in an electronic management system.

During this process, we prepared with others a book entitled *Effectiveness and Satisfaction in Antenatal Care* (Enkin and Chalmers 1982). Although this book was not the formal review that had been envisaged in 1976 (it contained only one meta-analysis [Grant and Mohide 1982]), it emphasized the importance of randomization for controlling selection biases when assessing the effects of care, and it drew heavily on the results of the RCTs of care during pregnancy that had been identified by the time of publication.

Searching for and classifying controlled trials reported over a period of nearly five decades took us more than five years. By 1985, we were able to publish a classified bibliography of more than 3,000 trials published between 1940 and 1984 (National Perinatal Epidemiology Unit 1985), and we established a systematic hand search of relevant journals (which currently number about 70) on an ongoing basis, so that reports could be identified promptly after publication.

In addition, we collaborated with others to seek details of unpublished perinatal trials by surveying 42,000 obstetricians and pediatricians in the 18 countries where the vast majority of controlled trials in perinatal medicine have been conducted (Hetherington et al. 1989). This survey was conducted in an attempt to address the problems presented by publication bias and its potentially adverse effect on the validity of reviews (Dickersin 1990). Unfortunately, the problems inherent in trying to identify unpublished trials are far from solved. Although there are signs of a developing willingness to improve the situation by registering trials at inception (*Lancet* 1991), reviewers remain heavily dependent on search strategies that are likely to result in incomplete identification of relevant unpublished studies (Chalmers 1990).

Commissioning and Conducting Systematic Reviews of the Effects of Care during Pregnancy and Childbirth

The first systematic review of the results of randomized trials using meta-analysis and involving care during pregnancy or childbirth was presented at an international congress in 1978, and published the following year (Chalmers 1979). The overview was based on published and unpublished data derived from four randomized trials that compared different

methods of monitoring the fetus during labor. The meta-analysis revealed a previously undetected distribution—unlikely to have occurred by chance—of 13 cases of neonatal seizures among the 2,000 or so babies who had been entered into the four trials. This observation generated the hypothesis that, compared with intermittent auscultation of the fetal heart, continuous fetal heart monitoring with an option to assess fetal acid-base status using scalp blood sampling reduced the risk of neonatal seizures. Our respect for the potential power of meta-analysis was strengthened when this hypothesis was subsequently tested and sustained in a randomized trial (for which one of us was a coinvestigator) that involved nearly seven times as many participants as the total number of women and babies in the previous four trials taken together (MacDonald et al. 1985).

It was against this background that, in 1986, after another sabbatical year's leave had become available to one of us (ME), we commissioned the systematic review of controlled trials that had been conceived ten years previously. The registers of published and unpublished trials described above provided the starting point for these reviews. Reviewers who agreed to collaborate in the project were provided with listings of trials likely to be relevant to their areas of responsibility, as well as copies of any papers they had difficulty in obtaining. Trials in languages other than English were translated if they seemed likely to contain important information. In addition, contributors to the review were given editorial guidelines for assessing the methodological quality of the studies listed for them, for abstracting the results in a form suitable for presentation in a systematic review, and for obtaining any unpublished information that seemed likely to improve the validity of the review (Chalmers et al. 1989b).

Data assembled in this way were entered in a centrally organized, electronic management system using specially commissioned software. These data formed the basis of systematic reviews upon which the texts of about half of the book's chapters were based. When the results of only one trial were available to assess the effects of a particular form of care, they were presented in the review (it is primarily for this reason that we have tended to refer to the analyses within the project as "systematic reviews" rather than "meta-analyses"). The effects of the many forms of care during pregnancy and childbirth that had not been evaluated using randomized trials were assessed using analyses of observational data. The

quality of the evidence on which the conclusions were based was described in each section.

Eventually, a 1500-page, two-volume book entitled *Effective Care in Pregnancy and Childbirth* (Chalmers, Enkin, and Keirse 1989a), containing hundreds of systematic reviews, was published. So that the principal conclusions of this large and expensive book (abbreviated *ECPC*) were accessible to women using the maternity services and to others, the findings were summarized in a 400-page, concurrently published paperback entitled *A Guide to Effective Care in Pregnancy and Childbirth* (*GEPC*) (Enkin, Keirse, and Chalmers 1989). Both books end with four appendices listing, respectively, forms of care (1) that have been shown to reduce the risk of negative outcomes; (2) that appear promising but require further evaluation; (3) with unknown effects; and (4) that are so unlikely to have beneficial effects that they should be abandoned. These appendices thus summarize evidence of relevance both for clinical practice and for deciding clinical research priorities.

Updating and Amending Systematic Reviews of Controlled Trials in the Light of New Data and Criticisms

In both books we warned our readers that, although we and our collaborators (within the resources available to us and to them) had tried to minimize bias in our analyses, we were aware that many of them might be improved (Chalmers et al. 1989b). We noted that these improvements might be effected by incorporating data that had not been available to us; by reanalyzing the information to which we had had access, using more thorough blinding during the selection of studies and abstraction of data; by using alternative aggregations of studies; possibly, by using one or other of the alternative statistical methods available; or by some combination of these steps. We urged others to conduct alternative analyses, using alternative materials and methods, so that the stability of the conclusions presented in the books could be assessed, and we undertook to ensure that improved and updated analyses, as well as new analyses, would be published promptly in electronic form. Between 1989 and 1992 these updated analyses were published in biannual issues of *The Oxford Database of Perinatal Trials* (*ODPT*) (Chalmers 1992).

In any event, no major mistakes have so far been drawn to our atten-

tion during the four years since the books were published. No form of care that we deemed capable of reducing some negative outcome of pregnancy has been shown not to do so; and no evidence has emerged to justify the retention of forms of care that we recommended abandoning. New evidence has, however, altered the status of some of the interventions whose effects were not clear from the evidence available at the time the book went to press. For example, it has become clear that periconceptional supplementation with folic acid in women who have previously given birth to babies with neural tube malformations (Lumley 1993), and prophylactic surfactant given soon after birth to infants at high risk of respiratory distress syndrome (Soll 1993) are both effective in reducing life-threatening neonatal morbidity. Whereas there was no strong evidence that any form of biophysical fetal assessment was useful at the time that the book went to press, subsequent evidence suggests that Doppler ultrasound assessment of umbilical artery waveforms in high-risk pregnancies should now be categorized as a promising technology meriting further investigation (Neilson 1993).

Appendix 2 summarizes the process of incorporating new data in the central database. This shows how trials are sought and classified, entered in the management system, and then distributed to one of about 30 reviewers. These reviewers, supported by a core editorial team, are responsible for updating the systematic reviews published in *ECPC*, and for preparing new reviews as new evidence emerges (Chalmers 1991). Following the recommendation of Mulrow, Thacker, and Pugh (1988), each review is presented in a structured report, a format that has facilitated the preparation of printed reviews, in the form both of newsletters published concurrently with every disk issue of *ODPT* (Chalmers 1992) and of leading articles commissioned for publication in the *British Journal of Obstetrics and Gynaecology* (Hofmeyr 1991, 1992; Carroli 1991; Lumley 1991). These reports will also help to smooth preparations for future publications using a variety of media. Appendix 3 shows the status of the overviews and records for which each member of the collaborative review group is responsible, as of October 1992.

Some Reflections on the Production of ECPC and ODPT

Frederick Mosteller, one of the commentators at the meeting organized by the International Society of Technology Assessment in Health Care to

consider the implications of our work, called for more information about the development and organization of the review process that led to publication of *ECPC* and *ODPT* (F. Mosteller 1992: personal communication). In a letter to one of us he said:

You had an army of authors and you must have had dedicated assistants. However friendly one is, it must have been a bit of a struggle to get everyone to do their meta-analyses in the same way. The turn-arounds required in such work (group writing) have, in my experience, been large in number and terribly time-consuming. How fast were things turned around? How many versions of manuscripts were created or revised? Were some items finally lost because you couldn't get them finished? How did you finance postage for the survey? These are not matters of idle curiosity (though I am curious); group projects are hard to run, and ideas for what makes things work, even if not backed by careful experimentation, may be worth a good deal.

Although we have little systematic data to offer in response to this request, we do have a number of "ideas for what will make things work"; however, we must stress that these views inevitably reflect only our own editorial perspective.

Esprit de Corps. By far the most important single reason for the success of these projects has been that the participants believe themselves to be engaged in an enterprise that can improve the care of women and babies during pregnancy and childbirth. This is not a sentimental opinion. It is a view that has often been reiterated to us by those who have helped the project to succeed, be they authors, computer programmers, clerks, secretaries, funding bodies, or our long-suffering families. This esprit de corps has led many people to contribute far beyond the call of any "duty" to the project, and many of these contributors have explicitly described the pride that their involvement in this research has brought them. They see the enthusiastic reception of the results by a wide variety of commentators because the research has helped to clarify which forms of care during pregnancy and childbirth do more good than harm, and vice versa.

Commissioning Reviews for ECPC. Because of the systematic approach that we wished each of our contributors to bring to the review process, we were asking a great deal more from each of them than is usually expected of people who are asked to write a chapter for a book (see

appendices in Chalmers et al. 1989b). Ideally, each of them was to combine knowledge about the topic he or she was being asked to review with a willingness to apply methodological rigor to the review process.

Sometimes we were fortunate to find that this combination of qualities existed within a single individual. Sometimes we arranged coauthorship “marriages” to try to ensure that content expertise and methodological rigor were applied to a particular topic. Sometimes we failed utterly. These failures were usually the result of knowledgeable people (usually very senior obstetricians) submitting manuscripts that revealed our failure to convey to them the nature of the “new kind of review” we were commissioning. Occasionally, the opposite situation caused problems; that is, methodological rigor was applied to the review without sufficient understanding of the material under analysis. As editors, we had to try to recognize and to correct these deficiencies.

We imposed no word limit on those whom we invited to contribute chapters to *ECPC*. Contributors were encouraged to minimize their discussion of the physiology, pathology, and epidemiology of the topic in question, but to use as many words as they felt were required to achieve adequate coverage of the effects of prophylactic and therapeutic strategies.

Editing ECPC. Each of us brought something different to the editorial task. One of us (ME) is an obstetrician with 35 years’ experience of mainly primary and secondary obstetric care. The experience of another (MJNCK) is mainly in secondary and tertiary obstetric care. The third member of the editorial team (IC) has only rudimentary clinical experience, but was able to contribute methodological expertise. This combination of attributes within the editorial team certainly generated a good deal of creative tension during the editing process; it also helped to ensure that specious arguments, poor logic, and egregious errors were detected more efficiently. We made a conscious effort to purge the book of unnecessary jargon and to make it “woman and baby” centered—a feature that has received favorable comments from a number of reviewers.

We asked contributors to the book who were able to do so to submit their manuscripts in electronic form as well as on paper. All the manuscripts were converted into the same word-processable form, which made the editorial task less daunting than it might otherwise have been. The manuscript was delivered to the publisher both in electronic form and on paper.

In spite of these technical aids, our editing task was substantial. It was

not unusual for chapters to go through ten or more drafts. Not infrequently, “editing” meant that we had to conduct the relevant review ourselves, from scratch, then rewrite the text completely in the light of the meta-analyses. Occasionally, the author(s) from whom the review had been commissioned were dropped completely (after we paid them a fee for their manuscript and promised them a complimentary copy of the published book). More usually, particularly if it was clear that the author(s) had tried hard to do a thorough job, the name of the editor who had done the most work was added as a coauthor of the chapter. In two or three instances, we abandoned plans to include chapters on particular topics, either because we decided that the topics were less relevant than we had originally envisaged, or because the available evidence was weak, or because we could identify no suitable author who would be likely to complete the task in the time available (we delivered the book to the publishers about 18 months later than we had intended).

Updating the Database and Preparing for Future Publications. The experience of editing *ECPC* has been invaluable in helping us to choose the team that is now responsible for updating the structured reports that constitute the database. Some of the people who authored *ECPC* chapters containing reviews of controlled trials have been dropped, and others (who have demonstrated their reliability since the book went to press) now have a central role in the project. In addition, we have involved a number of midwife researchers, one of them as a coeditor with us, all of whom are keen to contribute to this work.

If we had the opportunity to begin again, we might well have decided that, before attempting to prepare either of the books, all of the systematic reviews should have been completed, with the structured reports prepared and held in electronic form. We have adopted this approach in preparing for future publications, and it is one that we would urge others to consider if they are contemplating embarking on a similar exercise.

Funding the Preparation of ECPC and ODPT. We hope that the tangible results of our work may have made it easier for others to obtain funds for this kind of work than it was for us. As it happens, the Milbank Memorial Fund provided the first rebuff (in 1977) to our attempts to finance a start of the project! (Archie Cochrane and one of us [IC] had applied for funds to conduct a systematic hand search for reports of controlled trials in perinatal medicine.) Eventually the Maternal and Child Health Unit at WHO (using funds donated by the Swedish aid agency

SAREC) provided (between 1978 and 1985) the modest resources needed to organize the systematic literature search on which the project was founded, and to begin computerizing the database. Since then, specific project funds (to search for unpublished trials, to develop software, to provide secretarial and clerical support, and to prepare the paperback summary of the review) have been derived from a number of sources (listed in the Acknowledgments). It must be said, however, that the funds derived through these sources have covered only a small proportion of the costs of the work.

It would certainly have been difficult (and maybe impossible) to support the work we have described had it not been organized from within a multidisciplinary health services research unit (the National Perinatal Epidemiology Unit in Oxford), in which some scientific, programming, and clerical/secretarial staff and some nonstaff resources (space, equipment, consumables, telephone, postage, and so on) were "core funded" over a reasonable (four to six years) time scale. This form of support, which was provided by the Department of Health, and the facilities made available by some of the institutions with which editors and collaborators were associated, was, we believe, essential.

Discussion

To what extent can *ECPC* be considered a model for others? A companion volume reviewing randomized trials of neonatal care, *Effective Care of the Newborn Infant*, has already been prepared by Jack Sinclair, Michael Bracken, and their colleagues (Sinclair and Bracken 1992), and preparations are currently underway to keep its reviews up to date electronically. Although a working system now exists through which systematic reviews of RCTs of perinatal care can be prepared and updated as new evidence becomes available, considerable scope exists for improving the validity and efficiency of this process. Predictably, the availability of resources has been one of the constraints impeding these improvements.

It is no accident that the most impressive examples of systematic, up-to-date reviews are concerned with the prevention and treatment of cancer and cardiovascular disease. In addition to the substantial amounts of public funds and the large numbers of researchers available in these

fields, commercial and charitable funds are also plentiful relative to other areas of clinical research. Three examples of such reviews in these fields are particularly impressive: the Early Breast Cancer Trialists' Collaborative Group (1990); the Advanced Ovarian Cancer Trialists' Group (1991); and the Antiplatelet Trialists' Collaboration (1988). All known trialists in these three subject areas have been invited to contribute their data to centrally coordinated meta-analyses. Although these global collaborative enterprises began by using data referring to groups of patients, the reviews are increasingly based, as far as possible, on individual patient data. This facilitates quality control and enables reviewers to conduct analyses that depend on having longitudinal data on individual patients. Although resource intensive, this approach to systematic reviews is likely to become more widespread because of its obvious scientific merits.

If information of the kind that we have described is to become available to decision makers in health care and research, clinical scientists and funding agencies will have to invest increasingly in improving the scientific quality of reviews. Experience of organizing systematic, up-to-date reviews of controlled trials of health care remains limited. This was recognized in the United Kingdom when the recently established Research and Development Program of the National Health Service funded a center to facilitate this work. The opening of the center in November 1992 was welcomed all over the world. As a result, a substantial international collaborative effort—the Cochrane Collaboration—is now underway to ensure that systematic reviews of RCTs of health care will be prepared, maintained, and disseminated efficiently in future.

The database of reviews of RCTs in pregnancy and childbirth that we have described is the first of the modules of edited reviews to be contributed to the Collaboration's database: the Cochrane Database of Systematic Reviews. Electronic publication (on disk and online) began in May 1993. The collaborative review group responsible for maintaining the pregnancy and childbirth module of edited reviews is acting as one of the key test beds for the development and evaluation of standard methods for possible adoption by all of the Cochrane Collaboration's review groups. The evolution of other collaborative review groups will not follow exactly the same pattern as the one we have described; nevertheless, we hope this description of, and reflection on, our own experience may be of interest to others. Extensive worldwide collaboration is required to

accomplish the task of synthesizing existing information about the effects of health care efficiently. And efficiency is essential. People using the health services have already waited too long for the available evidence to be assembled and kept up to date.

References

- Acheson, D., and A.A.B., Poole. 1991. *Folic Acid in the Prevention of Neural Tube Defects* (PL/CMO(91)11/PL/CNO(91)6). London: Department of Health.
- Advanced Ovarian Cancer Trialists' Group. 1991. Chemotherapy in Advanced Ovarian Cancer: An Overview of Randomised Clinical Trials. *British Medical Journal* 303:884-93.
- Antiplatelet Trialists' Collaboration. 1988. Secondary Prevention of Vascular Disease by Prolonged Antiplatelet Treatment. *British Medical Journal* 296:320-31.
- Antman, E.M., J. Lau, B. Kupelnick, F. Mosteller, and T.C. Chalmers. 1992. Comparison of Results of Meta-Analyses of Randomized Control Trials and Recommendation of Clinical Experts: Treatments for Myocardial Infarction. *Journal of the American Medical Association* 268:240-8.
- Caroli, G. 1991. Management of Retained Placenta by Umbilical Vein Injection. *British Journal of Obstetrics and Gynaecology* 98:348-50.
- Chalmers, I. 1979. Randomised Controlled Trials of Fetal Monitoring 1973-1977. In *Perinatal Medicine*, eds. O. Thalhammer, K. Baumgarten, and A. Pollak, 260-5. Stuttgart: Thieme.
- . 1989. Evaluating the Effects of Perinatal Care. In *Effective Care in Pregnancy and Childbirth*, eds. I. Chalmers, M. Enkin, and M.J.N.C. Keirse, 3-38. Oxford: Oxford University Press.
- . 1990. Under-Reporting Research Is Scientific Misconduct. *Journal of the American Medical Association* 263:1405-8.
- . (1991). Improving the Quality and Dissemination of Reviews of Clinical Research. In *The Future of Medical Journals: In Commemoration of 150 Years of the British Medical Journal*, ed. S. Lock, 127-46. London: British Medical Journal.
- . Ed. 1992. *Oxford Database of Perinatal Trials*. Version 1.2, disk issue 7 (spring). Oxford: Oxford University Press.
- Chalmers, I., J. Hetherington, M. Newdick, et al. 1986. The Oxford Database of Perinatal Trials: Developing a Register of Published Reports of Controlled Trials. *Controlled Clinical Trials* 7:306-24.

- Chalmers, I., M. Enkin, and M.J.N.C. Keirse. Eds. 1989a. *Effective Care in Pregnancy and Childbirth*. Oxford: Oxford University Press.
- Chalmers, I., J. Hetherington, D. Elbourne, M.J.N.C. Keirse, and M. Enkin. 1989b. Materials and Methods Used in Synthesizing Evidence to Evaluate the Effects of Care during Pregnancy and Childbirth. In *Effective Care in Pregnancy and Childbirth*, eds. I. Chalmers, M. Enkin, and M.J.N.C. Keirse, 39–65. Oxford: Oxford University Press.
- Cochrane, A.L. 1972. *Effectiveness and Efficiency: Random Reflections on Health Services*. London: Nuffield Provincial Hospitals Trust.
- Collins, R., R. Gray, J. Godwin, and R. Peto. 1987. Avoidance of Large Biases in the Assessment of Moderate Treatment Effects: The Need for Systematic Overviews. *Statistics in Medicine* 6:245–50.
- Crowley, P., I. Chalmers, and M.J.N.C. Keirse. 1990. The Effects of Corticosteroid Administration before Preterm Delivery: An Overview of the Evidence from Controlled Trials. *British Journal of Obstetrics and Gynaecology* 97:11–25.
- Crowther, C.A. 1991. Hospitalisation for Bed Rest in Twin Pregnancy. In *Oxford Database of Perinatal Trials*, ed. I. Chalmers, 3376. Version 1.2, disk issue 6 (autumn).
- Dickersin, K. 1990. The Existence of Publication Bias and Risk Factors for its Occurrence. *Journal of the American Medical Association* 263:1385–9.
- Dickersin, K., P. Hewitt, L. Mutch, I. Chalmers, and T.C. Chalmers. 1985. Perusing the Literature. Comparison of MEDLINE Searching with a Perinatal Trials Database. *Controlled Clinical Trials* 6:306–17.
- Dickersin, K., K. Higgins, and C.L. Meinert. 1990. Identification of Meta-Analyses: The Need for Standard Terminology. *Controlled Clinical Trials* 11:52–66.
- Doll, R. 1991. Development of Controlled Trials in Preventive and Therapeutic Medicine. *Journal of Biosocial Science* 23:365–78.
- Early Breast Cancer Trialists' Collaborative Group. 1990. *Treatment of Early Breast Cancer*. Volume 1: *Worldwide Evidence 1985–1990*. Oxford: Oxford University Press.
- Enkin, M., and I. Chalmers. 1982. *Effectiveness and Satisfaction in Antenatal Care* (Clinics in Developmental Medicine nos. 81/82). London: Spastics International Medical Publications/William Heinemann Medical Books.
- Enkin, M., M.J.N.C. Keirse, and I. Chalmers. 1989. *A Guide to Effective Care during Pregnancy and Childbirth*. Oxford: Oxford University Press.
- Grant, A., and P. Mohide. 1982. Screening and Diagnosis in Antenatal

- Care. In *Effectiveness and Satisfaction in Antenatal Care*, eds. M. Enkin and I. Chalmers, 22–59 (Clinics in Developmental Medicine nos. 81/82). London: Spastics International Medical Publications/William Heinemann Medical Books.
- Grant, A., J. Sleep, H. Ashurst, and J.A.D. Spencer. 1989. Dyspareunia associated with the use of Glycerol-impregnated Catgut to Repair Perineal Trauma—Report of a Three-Year Follow-up Study. *British Journal of Obstetrics and Gynaecology* 96:741–3.
- Hetherington, J., K. Dickersin, I. Chalmers, and C.L. Meinert. 1989. Retrospective and Prospective Identification of Unpublished Controlled Trials: Lessons from a Survey of Obstetricians and Pediatricians. *Pediatrics* 84:374–80.
- Hofmeyr, G. J. 1991. External Cephalic Version at Term: How High are the Stakes? *British Journal of Obstetrics and Gynaecology* 98:1–7.
- . 1992. Amnioinfusion: A Question of Benefits and Risks. *British Journal of Obstetrics and Gynaecology* 99:449–51.
- International Register of Perinatal Trials*. 1991. Department of Pediatrics, University of Vermont, and National Perinatal Epidemiology Unit, Oxford.
- ISIS-2 Collaborative Group. 1988. Randomised Trial of Intravenous Streptokinase, Oral Aspirin, Both, or Neither among 17,187 Cases of Suspected Acute Myocardial Infarction: ISIS-2. *Lancet* 2:349–60.
- Lancet*. 1991. Making Clinical Trialists Register. *Lancet* 338:244–5.
- Lumley, J. 1991. Stopping Smoking—Again. *British Journal of Obstetrics and Gynaecology* 98:847–52.
- . 1993. The Effect of Periconceptional Folate (4 mg/day) vs. Placebo on the Recurrence of Neural Tube Defects. In *Pregnancy and Childbirth Module*, eds. M.W. Enkin, M.J.N.C. Keirse, M.J. Renfrew, and J.P. Neilson. Cochrane Database of Systematic Reviews: Review No. 06488, April 2, 1992. Published through Cochrane Updates on Disk, Oxford: Update Software Ltd.
- MacDonald, D., A. Grant, M. Sheridan-Pereira, P. Boylan, and I. Chalmers. 1985. The Dublin Randomised Controlled Trial of Intrapartum Fetal Heart Rate Monitoring. *American Journal of Obstetrics and Gynecology* 152:524–39.
- MacMahon, S., R. Collins, R. Peto, R.W. Koster, and S. Yusuf. 1988. Effects of Prophylactic Lidocaine in Suspected Acute Myocardial Infarction: An Overview of Results from the Randomized, Controlled Trials. *Journal of the American Medical Association* 260:1910–16.
- Medical Research Council Vitamin Study Research Group. 1991. Prevention of Neural Tube Defects: Results of the Medical Research Council Vitamin Study. *Lancet* 338:131–7.
- Mugford, M., J. Piercy, and I. Chalmers. 1991. Cost Implications of Dif-

- ferent Approaches to the Prevention of Respiratory Distress Syndrome. *Archives of Diseases in Childhood* 66:757-64.
- Mulrow, C.D. 1987. The Medical Review Article: State of the Science. *Annals of Internal Medicine* 104:485-8.
- Mulrow, C.D., S.B. Thacker, and J.A. Pugh. 1988. A Proposal for More Informative Abstracts of Review Articles. *Annals of Internal Medicine* 108:613-15.
- National Perinatal Epidemiology Unit. 1985. *A Classified Bibliography of Controlled Trials in Perinatal Medicine 1940-1984*. Oxford: Oxford University Press.
- Neilson, J.P. 1993. Doppler Ultrasound in High Risk Pregnancies. In *Pregnancy and Childbirth Module*, eds. M.W. Enkin, M.J.N.C. Keirse, M.J. Renfrew, and J.P. Neilson. Cochrane Database of Systematic Reviews: Review No. 03889, January 19, 1993. Published through Cochrane Updates on Disk, Oxford: Update Software Ltd.
- Peto, R. 1987. Why Do We Need Systematic Overviews of Randomized Trials? *Statistics in Medicine* 6:233-40.
- Sinclair, J.C., and M. Bracken. Eds. 1992. *Effective Care of the Newborn Infant*. Oxford: Oxford University Press.
- Soll, R.F. 1993. Prophylactic Administration of Any Surfactant. In *Pregnancy and Childbirth Module*, eds. M.W. Enkin, M.J.N.C. Keirse, M.J. Renfrew, and J.P. Neilson. Cochrane Database of Systematic Reviews: Review No. 05664, March 24, 1993. Published through Cochrane Updates on Disk, Oxford: Update Software Ltd.
- Williamson, J.W., P.S. German, R. Weiss, E.A. Skinner, and F. Bowes. 1989. Health Science Information Management and Continuing Education of Physicians. *Annals of Internal Medicine* 110:151-60.

Acknowledgments: The work described in this article has involved collaboration involving literally hundreds of people. Most of them have been acknowledged by name in *Effective Care in Pregnancy and Childbirth*; we reiterate our indebtedness to all of them here, and acknowledge our particular debts to Eleanor Enkin, Jini Hetherington, Sally Hunt, and Mark Starr. The main sources of funds for the work have been the Department of Health (London), the World Health Organization, Oxford University Press, the Rockefeller Foundation, and the National Center for Health Services Research. We are also grateful for helpful comments on earlier drafts of this article provided by Jonathan Lomas, Frederick Mosteller, Jack Sinclair, Jane Sisk, Barbara Stocking, Elizabeth Wilson, and colleagues at the National Perinatal Epidemiology Unit, and to Caroline Caldicott for help in preparing the manuscript.

Address correspondence to: Iain Chalmers, the UK Cochrane Centre, NHS R&D Programme, Summertown Pavilion, Middle Way, Oxford OX2 7LG, United Kingdom.

Appendix 1

Chronology of Steps Taken to Establish a Continuously Updated Review of the Effects of Perinatal Care

Year	Event
1972	Importance of randomized trials communicated effectively by Archie Cochrane in his monograph, <i>Effectiveness and Efficiency: Random Reflections on Health Services</i>
1974	Card file of references to controlled trials in perinatal medicine established; MEDLINE search strategy designed and implemented monthly
1976	Outline plan formulated for a systematic review of all perinatal trials, including overviews of similar trials to reduce random error in estimating effects of care
1977	Unsuccessful attempts made, in collaboration with Archie Cochrane, to obtain funds for a hand search of 60 "core" journals back to 1950
1978	Grant provided by Maternal and Child Health Unit, WHO, Geneva, enabling systematic hand search for reports of perinatal trials to begin
1979	First overview (meta-analysis) of perinatal trials published
1980	Introduction of pilot classification system for perinatal trials
1981	Distribution of provisional listing of controlled trials to identify important omissions Letter to <i>Lancet</i> calling for registration of perinatal trials
1982	Microcomputer funded by WHO for storage of information about registered trials in a database Publication of a book by Enkin and Chalmers (<i>Effectiveness and Satisfaction in Antenatal Care</i>) reviewing controlled trials of antenatal care, but without using meta-analysis
1983	Development of software for more flexibly manipulating the database search of perinatal trials
1984	Implementation of amended classification system for perinatal trials and coding of more than 3,000 trials
1985	Publication of <i>Classified Bibliography of Controlled Trials in Perinatal Medicine 1940-1984</i> in book form (National Perinatal Epidemiology Unit 1985)
1985	Comparison of MEDLINE with database of perinatal trials reveals 50 percent shortfall in identification using MEDLINE Grant from Oxford University Press to develop database for eventual release as an electronic publication

- 1986 Development of database of perinatal trials documented in the journals *Controlled Clinical Trials* and *WHO Chronicle*
Systematic reviews commissioned of all controlled trials (about 3,500) in obstetrics
- 1987 Software of database of perinatal trials “beta-tested”
Letters sent to 42,000 clinicians in 20 countries seeking registration of unpublished trials
Reviews and overviews commissioned of all controlled trials (about 1,500) in neonatal pediatrics
- 1988 Publication of *Oxford Database of Perinatal Trials (ODPT)* (version 1.0, disk issue 1) (I. Chalmers, ed.)
Publication of the first in a series of overviews (meta-analyses) in the *British Journal of Obstetrics and Gynaecology*
- 1989 Publication of *Effective Care in Pregnancy and Childbirth (ECPC)* (Chalmers, Enkin, and Keirse, eds.)
Publication of *A Guide to Effective Care in Pregnancy and Childbirth (GECPC)* (Enkin, Keirse, and Chalmers, eds.)
Grant from National Institute for Child Health and Human Development to assess feasibility of prospective registration of perinatal trials at inception/in progress
ODPT disk issues distributed biannually
- 1990 Development and introduction of structured reports of systematic reviews published within *ODPT*
Publication of the first in a series of commentaries for the *British Journal of Obstetrics and Gynaecology* based on systematic reviews published electronically
- 1991 Registration of team of obstetric and midwifery reviewers for RCTs in pregnancy and childbirth
Introduction of four-page newsletters published with each biannual disk issue of *ODPT*
Publication of an account of *ECPC* and *ODPT* by Chalmers in *The Future of Medical Journals* (a book marking 150 years of the *British Medical Journal*)
- 1992 Publication of *Effective Care of the Newborn Infant (ECNI)* by Sinclair and Bracken
Final disk issue of *ODPT* published
Collaborative Review Group for Pregnancy and Childbirth designated piloting/demonstration project for the Cochrane Collaboration

- 1993 Systematic reviews of RCTs in pregnancy and childbirth contributed as a module to the Cochrane Database of Systematic Reviews, and disseminated on disk and online
-

Appendix 2

Steps in Maintaining a Continuously Updated Review of the Effects of Perinatal Care

1. Identify potentially eligible studies
 - *Published* reports of trials:
 - ongoing hand search of 70 "core" journals
 - ongoing monthly MEDLINE search
 - ongoing "chance discovery"
 - *Unpublished* trials:
 - ad hoc survey of 42,000 practitioners (1987-88)
 - ad hoc survey of funding bodies/academic departments (1990-91)
 - ongoing "chance notification"
2. Obtain copy of the report, check allocation "random" or "quasi-random" and intervention "perinatal"
 - Write to authors for clarification, if necessary
3. Enter eligible trials on central database
 - Accession details
 - Full bibliographic reference
 - Number and characteristics of participants
 - Number and characteristics of intervention(s)
 - Outcomes reported
 - Country of trial
 - Member of review group responsible
4. Distribute topic-specific work lists to reviewers
 - Titles of existing reviews
 - Records already coded "hold," pending possible incorporation in an overview
 - Records requiring "status" codes
 - Trials registered as unpublished
 - Trials registered as "recruiting" or "planned"
5. Describe tasks to reviewers
 - Assign status to each trial record
 - Contact authors for clarifications and/or additional data
 - Assess methodological quality of trials
 - Tabulate characteristics of trials potentially eligible for inclusion in an overview

- Abstract data for incorporation (if appropriate) in an existing or new systematic review
 - Conduct overview analyses using standard software
 - Write structured report of systematic review
 - Identify amended systematic reviews in which substantive changes have been made
 - Submit new and amended overviews (and updated worklists) for incorporation in central database
6. Update the core database
 - Central editorial checking of new and amended reviews
 - Assign new and amended status codes
 - Incorporate new and amended reviews
 7. Prepare data for dissemination
 - Online electronic publication
 - Disks
 - Newsletters
 - Books
-

Appendix 3

Summary of Reviewer Status of Pregnancy and Childbirth Reviewers within the Cochrane Centre Pregnancy and Childbirth Module (as of October 1992) (*Appendix 3 table follows on next four pages.*)

APPENDIX

Editor(s)	Topics covered	Total no. of overviews
Collins, Rory [R1.10]	Hypertension, prevention	2
	Hypertension, treatment	8
Crowley, Patricia [R1.16]	Fetal maturation	6
	PROM preterm (steroids/antibiotics)	2
	Postterm pregnancy	4
Crowther, Caroline [R1.07]	Multiple pregnancy	6
	Bed rest for hypertension	3
	Rhesus	5
Duley, Lelia [R1.20]	Hypertension, prevention	3
	Hypertension, treatment	23
Elbourne, Diana [R1.05]	Third stage of labor	12
Enkin, Murray [R1.02]	Cesarean section (miscellaneous)	13
Fraser, Bill [R1.28]	Dystocia	7
Grant, Adrian [R1.04]	Prenatal diagnosis	12
	Cervical cerclage	7
	Fetal therapy/preterm delivery	4
	Intrapartum fetal monitoring/therapy	12
	Repair of perineal trauma	10
Hannah, Mary [R1.23]	PROM at term	6
Hodnett, Ellen [R1.11]	Social support during pregnancy	2
	Social support during labor	2
	Other social support	0
Hofmeyr, Justus [R1.08]	Fetal malposition	7
	Conservative management of compromised fetus in pregnancy and labor	10
		7
Howell, Charlotte [R1.19]	Pain relief in labor	10
	Hydration/antacids	10
Hunter, David [R1.24]	Diabetes	6
	Gestational diabetes	4
Jewell, David	Gastrointestinal symptoms in pregnancy	12
Young, Gavin [R1.22]	Vaginitis	4
	Other symptoms (in pregnancy)	8
Johanson, Richard [R1.13]	Instrumental delivery	7

3

No. of overviews needing commentaries	No. of records with no "S" code	No. of records coded "hold"	Total no. of records in 1991	Total no. of records	Date last contact received	Date last contact sent	Notes
1	43	10	9	116	1/9/92	1/7/92	
2							
0	12	1	8	120	1/17/92	1/17/92	
0							
0	6	0	1	25	7/14/92	7/28/92	
0							
5							
0	45	11	16	159	4/27/92	6/2/92	
2							
5	9	23	9	101	7/18/92	2/28/92	
0	12	5	12	156			
2	10	6	7	35	7/17/92	7/9/92	
3	61	2	13	155		5/6/92	
0							
0							
0							
0							
2	12	2	8	40	6/21/92	7/8/92	
2	74	0	12	93	5/7/92	5/7/92	
2							
0							
0	13	0	10	86	4/1/92	4/3/92	
0							
0							
5	871	0	50	924	7/25/92	5/19/92	
9							
0	11	2	2	45	7/19/91	10/10/91	
0							
0	8	6	7	65	1/20/92	1/16/92	
1							
0							
0	3	0	6	35	7/25/92	8/3/92	

(continued)

APPENDIX

Editor(s)	Topics covered	Total no. of overviews
Kaufman, Karyn [R1.18]	Routine antenatal care	1
	Perineal care	17
Keirse, Marc [R1.01]	Preterm labor, prevention	7
	Preterm labor, treatment	18
	PROM preterm (contractility)	2
	Cervical ripening	37
	Induction	23
Kramer, Michael [R1.25]	Fetal death	3
	Energy/protein intake	5
Lumley, Judith [R1.21]	Maternal allergen avoidance/ atopic disease	3
	Work and physical activity in pregnancy	1
	Pre/periconceptual care	3
Mahomed, Kassam	Antismoking interventions	5
	Anemia, prevention	3
Hyttén, Frank [R1.06]	Anemia, treatment	0
	Sickle cell anemia/Pyridoxine/ vitamin D/zinc	5
	Other vitamins/minerals	1
	Bleeding in pregnancy	0
	Antimalarials	0
	Other tests of fetal well-being	11
Neilson, Jim [R1.09]	Imaging ultrasound	5
	Doppler ultrasound	2
	Other tests of fetal well-being	11
Prendiville, Walter [R1.26]	Recurrent and threatened miscarriage	8
Renfrew, Mary [R1.03]	Routine intrapartum care	5
	Routine puerperal care	5
	Lactation	14
	Lactation suppression	31
Smaill, Fiona [R1.12]	Urinary tract infection	14
	Antibiotics and cesarean section	16
	Group B streptococcus	3
	Postpartum infection	3
	Other infections	3
Spiby, Helen [R1.17]	Position during labor	4
	Pushing during second stage	2
Unassigned [R9.99]	Medical conditions	0

Abbreviation: PROM, prelabor rupture of membranes.

3 (continued)

No. of overviews needing commentaries	No. of records with no "S" code	No. of records coded "hold"	Total no. of records in 1991	Total no. of records	Date last contact received	Date last contact sent	Notes
0	17	6	3	65	7/13/92	7/16/92	
17							
2	175	16	67	745	7/15/92	7/8/92	
2							
0							
17							
12							
0							
1	9	0	0	93			
0							
0							
0	19	0	8	44	2/19/92	7/7/92	
0							
0	21	14	10	166	4/29/92	2/25/92	
0					(KM)	(KM)	
0						2/26/92	
0						(FH)	
0							
0							
0	59	0	8	164	7/28/2	7/22/92	
0							
0							
2	13	0	3	42	6/1/92	6/8/92	
2	140	0	14	352			
5							
13							
0							
0	66	12	7	150	7/14/92	7/28/92	
0							
0							
0							
0	23	4	8	80	8/24/92	8/28/92	
2							
0	11	0	1	7			