
The Cochrane Collaboration and the process of systematic reviewing

Clive Adams & Karla Soares

In 1979 Archie Cochrane, a British epidemiologist, stated that:

“it is surely a great criticism of our profession that we have not organised a critical summary, by specialty or subspecialty, adapted periodically, of all relevant randomized controlled trials.”

Cochrane, 1979

Cochrane said that, within the UK National Health Service, it was the tuberculosis physicians who were practising nearest to the best evidence of the time, while he gave the ‘wooden spoon’ to the obstetricians.

Dr Iain Chalmers, an epidemiologist and obstetrician, took up this challenge. Over a decade ago, he and his colleagues began to collect all randomised controlled trials relevant to the care of mothers and babies. They searched electronic databases, hand-searched 60 journals, and contacted 40 000 obstetricians across the world in order to identify unpublished material (Chalmers *et al*, 1986). Currently the Perinatal Register contains approximately 7000 controlled trials. This continually updated register is the basis for the 600 systematic reviews that have been electronically published for the past decade (Enkin *et al*, 1992). These reviews continue to be updated, every three months, as new evidence is identified (Enkin *et al*, 1993).

There have been many paper publications of these individual systematic reviews. They were also collected in two large academic volumes (Chalmers *et al*, 1989). In addition, a summary of all these findings was published for consumers of care (Enkin *et al*, 1996). The final chapter of this book outlines the treatments that are of proven worth, those that are not proven one way or another, and those that are clearly harmful.

This colossal task of systematic compilation of data on effectiveness, undertaken by Chalmers and his colleagues, became the pilot study for the Cochrane Collaboration, which is attempting to identify every randomised trial of health care and produce and maintain systematically conducted reviews of these studies (Chalmers *et al*, 1992).

The National Health Service Research and Development Initiative supported the founding of the first Cochrane Centre in the UK. There are now other centres in Africa, Australia, Brazil, Canada, Denmark, France, Italy, The Netherlands, and four in the USA, as well as reviewers in many more countries. The work of the Collaboration has even been likened (favourably) to that of the Human Genome Project (Naylor, 1995).

The UK Cochrane Centre facilitates the process by which health care groups become established, and helps develop, produce and maintain the Cochrane Library. This is an inexpensive electronic publication that is updated every three months. It is published by Update Software and, hopefully, will be available on the Internet in late 1997 (see Appendix).

The Cochrane Library contains several databases. The ‘Cochrane Controlled Trials Register’ holds references to approximately 120 000 randomised or quasi-randomised trials identified by those working within the Collaboration. The ‘Database of Reviews of Effectiveness’ contains reports and commentaries on about 1500 already paper-published systematic reviews and is assembled and maintained by the Centre for Reviews and Dissemination in York, UK. The ‘Cochrane Database of Systematic Reviews’ is the flagship of the Cochrane Library. It holds regularly maintained, systematically conducted reviews on every

Dr Clive Adams, psychiatrist and epidemiologist, works at the UK Cochrane Centre (Summertown Pavilion, Middle Way, Oxford OX2 7LG) and also coordinates the Cochrane Schizophrenia Group. His clinical work is based in the University Department of Psychiatry, Oxford. Dr Karla Soares, psychiatrist and epidemiologist from São Paulo, Brazil, works on reviews of the management of tardive dyskinesia, currently with the Cochrane Schizophrenia Group in Oxford.

aspect of health care. These reviews are produced not only for clinicians and researchers, but also for managers, policy makers and recipients of care.

The Cochrane Collaboration and mental health

Currently, there are two mental health groups working within the Collaboration, the longest established of which is the Cochrane Schizophrenia Group (Cochrane Collaboration Schizophrenia Review Group, 1994). This group is identifying and reviewing every intervention for those with any non-affective, non-organic psychotic illness or any chronic/severe mental illness where a clear diagnosis has not been specified. The Depression, Anxiety and Neurosis Group is focusing not only on the affective disorders, but also on eating disorders, somatisation problems and deliberate self-harm. The management of those with addictions, behavioural problems and learning disabilities is the focus of other groups in their early stages of genesis.

The Collaboration is still relatively young and much effort has been invested in organising the infrastructure by which reviewers can be supported. Currently there are only 11 systematic reviews in the Cochrane Library directly related to people with severe mental illnesses. These reviews deal with clozapine; ECT; family intervention for those with schizophrenia; vitamin E; γ -aminobutyric acid agonists; cholinergics and anticholinergics; benzodiazepines and calcium-channel blockers for the management of neuroleptic-induced tardive dyskinesia; antipsychotics for people with both schizophrenia and learning disabilities; and case management for severe mental illness.

Traditional *v.* systematic reviews

Traditionally, someone undertaking a review of a particular intervention would collect together information, summarise these data, usually qualitatively, and produce a report. Even if the search for included data had been thorough, the methods of the reviewing process itself frequently would not be explicit. In fact, 98% of treatment reviews within four major medical journals contained no 'Methods' section whatsoever and only 6% attempted quantitative analysis. Despite

this lack of rigour, conclusions were drawn in 74% of the reviews (Mulrow, 1987).

Powerful illustrations already exist of how the results and recommendations of reviews, conducted in a systematic way differ from those of more traditional commentaries. The widely quoted example comes from cardiology. In 1992, Antman *et al* undertook a cumulative systematic review, spanning 30 years, of various treatments used for people with acute myocardial infarction. They pooled the results of similar trials that involved comparable participants and interventions. For example, the first trial of thrombolytic therapy took place in the early 1960s but had so few participants (23) that a clear result favouring either thrombolytic therapy or the placebo was not possible. The next trial was a few years later and the cumulative total of participants began to increase. By the early 1970s the cumulative total was in the thousands and thrombolytic therapy was shown to be effective in reducing mortality ($P>0.01$). The very large trials undertaken thereafter only increased the degree of certainty that thrombolytic therapy is indicated for those with acute myocardial infarct (by 1990 $P>0.0001$). Antman *et al* (1992) also looked at what the key traditional reviews and texts were saying over the same three decades. It became clear that it took about a decade for traditional reviews to catch up with the best available summated evidence. Indeed, over a decade after a systematic review had shown otherwise, some key texts were omitting to recommend thrombolytic therapy. By this omission, traditional reviews were recommending a care regime that was, at best, harmful, and at worst, lethal.

A systematic review of trials

The systematic approach to reviewing makes the process of summarising literature more objective, open to scrutiny and valid criticism. A systematic review, like any retrospective survey of data, involves the employment of basic epidemiologic principles. Outlined below is an approach to undertaking a systematic review. A specific example is used to help illustrate some issues.

Scenario

A clinician/researcher and a colleague are invited, by hospital managers, to proffer an opinion on how implementing the Care Programme Approach (CPA; Department of Health, 1991) will affect their local mental health services. They decide to review,

systematically, all relevant trials to see whether there is a clear estimate of effect and what evidence exists for this care policy, now a statutory obligation within the UK.

The protocol (Box 1)

As with any well conducted survey, a protocol is written in order to help the reviewers think through what are the truly important issues and to pre-state the criteria by which data for the review are to be identified, selected, extracted and assimilated.

Background

This would justify the importance of this particular systematic review. The reviewers may wish to describe the basis of an hypothesis they are planning to investigate or summarise relevant information from non-trial reports. In our example, descriptive studies may have suggested that those with chronic illness seem to benefit more from CPA than those with a short history of illness. This finding could be stated in the background and then investigated within a sensitivity analysis.

Defining the criteria

The participants In the CPA example the reviewers may want to be very specific and review the effectiveness of CPA only for people with schizophrenia. On the other hand, they may decide that this is likely to be less generalisable – less like the real world – and choose to focus on a participant group that included anybody that was ‘chronically or seriously mentally ill’. It is feasible that the reviewers would stipulate that only those with ‘chronic/serious mental illnesses’, as defined by use of specific operational criteria, be included. Again, thinking pragmatically, operational criteria are rarely used in everyday practice for the people who receive the CPA and so the reviewers could well conclude that the use of operational criteria was not mandatory for data to be included. This will make the review more generalisable.

The interventions Continuing the example, the CPA is one term among a multitude of terms referring to attempts to provide community care programmes in some sort of managed way. The reviewers may wish to define the CPA as equivalent to low-intensity ‘case management’, as practised in other parts of the world. In turn, ‘low-intensity case management’ may be defined as, for example, a ratio of one case manager to at least 15 clients.

The reviewers will also want to define the comparison group. Although interested in the

Box 1. Steps of a Cochrane systematic review (see Sackett, 1994)

Step 1 – the protocol

Cover sheet
Background
Objectives
Materials
Criteria for considering trials for review – types of participants; types of intervention; types of outcome measures
Methods – search strategy for identification of studies; quality assessment; data extraction and management
Initial references
Acknowledgements

Step 2 – the review

Characteristics of included studies
Characteristics of excluded studies
Results
Discussion
Implications for practice
Implications for research
References
Abstract

effectiveness of assertive community treatment, where there are fewer clients to every carer than in low-intensity case management, it is the effectiveness of the CPA when compared with the standard care package that concerns the hospital managers.

The outcomes The reviewers have the luxury of being able to consider not so much what outcomes are likely to be available in the data, but, rather, what outcomes are truly of interest. For example, a research readership may be interested in any improvement on a given mental state scale, whereas clinicians may be interested only in what is a relevant shift on the same scale as seen from the perspectives of patients or their carers. Recipients of care, on the other hand, may not be interested at all that CPA facilitates a 20% shift in the Brief Psychiatric Rating Scale (Overall, 1962), a common outcome in mental health trials, but may want to know whether the intervention helped the person stay in employment, out of hospitals, or avoid self-harm or death. The managers of the hospital will, among other things, be interested in economic outcomes, although the researchers may realise that such data are rare within trials. In this example, the reviewers, by describing outcomes

readily reported in trials, as well as important outcomes that are not available, help to highlight gaps in current knowledge.

Identification of material

The reviewers will clearly stipulate how the relevant data were identified. Finding all relevant data, in order to minimise potential for the introduction of selection bias and the play of chance (random error; see Chalmers, 1989) is not straightforward. There are many problems with electronic searching. Not only can there be faulty software (Adams *et al*, 1992), inconsistent or poor-quality indexing (Dickersin *et al*, 1994; Hay & Adams, 1997), but searching in one language will fail to identify material in another (details available from author) and, probably most importantly, inexperience at undertaking searches results in large proportions of material being missed. For example, when trying to identify trials in 'high-yield' journals within Medline an experienced searcher found only 50% of relevant material, the other half being largely impossible to find by using methodology terms; however, a person with a normal level of experience of searching Medline found only 20% of the total (Adams *et al*, 1994). Because of the shortcomings of electronic databases, hand searches must be undertaken. This means turning over every single page of relevant journals in order to identify material which is or could be a randomised controlled trial. Many of the groups within the Cochrane Collaboration have already undertaken large amounts of hand searching. The Cochrane Schizophrenia Group, for example, has hand searched the *British Journal of Psychiatry*, *Archives of General Psychiatry*, *American Journal of Psychiatry*, *Acta Psychiatrica Scandinavica*, and *Psychological Medicine* from the first issue after 1948 to the present day, and all this material, not just trials relating to schizophrenia, is available within the Cochrane Library (Cochrane Controlled Trials Register).

In our example, if the reviewers searched electronic databases such as Embase, PsycLit or Medline (there is evidence that searching more than one is prudent), it should be made explicit exactly how the search was performed (see Box 2). The reviewers may wish to hand search high-yield literature such as specialist journals or conference proceedings, contact authors, or use citation searching to trace additional relevant trials.

Application of the selection criteria

By running a Medline search, such as that outlined in Box 2, many hundreds of citations will be identified. These will have to be inspected, and

material that is relevant acquired. This task must be done reliably so the reviewers may select studies independently. This will result in three groups of citations – those that are of agreed interest, those that are in dispute, and those that seem by both reviewers to be of no relevance to the review. The first two groups of articles would be acquired. In turn, the full reports will have to be inspected to see whether they meet the pre-specified inclusion criteria. Again, to make this process more reliable, the reviewers may want to work independently, inspect the full reports, resolve disputes verbally or, where this is impossible, add that particular study to those awaiting assessment while they contact the authors for more information. For example, if a study report mentions that, "the double-blind research method was used", it may not be entirely clear as to whether randomisation took place. Seeking information from the original authors may be the only way to resolve this dilemma.

Quality control

It is important that once reports have been selected and found to meet the inclusion criteria that some quality filter is applied. There is reasonable empirical evidence that the quality of the reporting of the randomisation procedure is an important indicator of the potential of the introduction of bias into the study. Chalmers *et al* (1983) showed that, as the quality of reporting of randomisation increased, the proportion of studies with statistically significant results, and the numbers of researchers recommending the specific intervention, decreased. This finding has since been replicated (Schulz *et al*, 1994). Other quality parameters, such as 'blindness' to allocation and 'intention-to-treat analysis', or at least reporting the reasons for people dropping out before the end of the study, are also important.

The reviewers must clearly stipulate where the quality cut-offs are going to be for trials and how these decisions regarding quality are to be made. Again, independent, reliable decisions are important. These decisions may be very pragmatic and include trials that merely state that the participants were randomised, or may be formalised by employing one of the many trial quality scales (Jadad, 1996).

Data management

Data extraction should be reliable and the reviewers may wish to follow the procedure described for the selection of trials. Once more, independent, reliable extraction, perhaps utilising a data extraction form, is vital, and additional

Box 2. Sample search strategy for a case management review (Medline specific)

Trial search of high specificity

- #1 explode RANDOMIZED-CONTROLLED-TRIALS in MeSH (all subheadings)
- #2 RANDOMIZED CONTROLLED TRIAL in PT
- #3 explode RANDOM-ALLOCATION in MeSH (all subheadings)
- #4 explode DOUBLE-BLIND-METHOD in MeSH (all subheadings)
- #5 explode SINGLE-BLIND-METHOD in MeSH (all subheadings)
- #6 explode CLINICAL-TRIALS in MeSH (all subheadings)
- #7 CLINICAL TRIAL in PT
- #8 CONTROLLED CLINICAL TRIAL in PT
- #9 RANDOMI*
- #10 RANDOMLY* near2 (ASSIGN* or ALLOCATE*)
- #11 CLIN* near TRIAL*
- #12 (SINGL* or DOUBL* or TREBL* or TRIPL*) near (BLIND* or MASK*)

Trial search of lower specificity

- #13 explode PLACEBOS in MeSH (all subheadings)
- #14 PLACEBO*
- #15 RANDOM*
- #16 explode RESEARCH-DESIGN in MeSH (all subheadings)
- #17 TG=COMPARATIVE-STUDY
- #18 explode EVALUATION-STUDIES in MeSH (all subheadings)
- #19 FOLLOW-UP-STUDIES in MeSH (all subheadings)
- #20 PROSPECTIVE-STUDIES in MeSH (all subheadings)

General chronic mental illness search

- #21 explode SCHIZOPHRENIA in MeSH (all subheadings)
- #22 explode PARANOID-DISORDERS in MeSH (all subheadings)
- #23 SCHIZO*
- #24 HEBEPHRENI*
- #25 OLIGOPHRENI*
- #26 (PSYCHOTIC* or PSYCHOSIS or PSYCHOSES) not NEUROTIC*
- #27 [(CHRONIC* or SEVER*) near2 MENTAL*] near2 (ILL* or DISORDER*)

Specific case management search

- #28 (CASE or CARE) and (MANAGEMENT)
- #29 COMMUNITY near1 TREATMENT
- #30 TCL
- #31 MADISON

Combined search

[(#1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13 or #14 or #15 or #16 or #17 or #18 or #19 or #20) and (#21 or #22 or #23 or #24 or #25 or #26 or #27) and (#28 or #29 or #30 or #31)]

Note: The use of operators ('or', 'near2' (within two words of), 'and') and Boolean logic facilitates search sensitivity and precision. A set of methodological phrases is connected by the word 'and' to a phrase specific for serious mental illness. Lastly, and again using the word 'and', this is connected to the specific case management search to form the combined search (TCL and Madison are synonyms for case management used in the USA)

information may have to be sought from the authors of the studies.

The systematic reviewers would pre-state how data are to be analysed. How are dichotomous data to be summated, how are con-

tinuous data to be managed, how are measurements of similar outcomes on differing scales to be dealt with and what is to be done with incomplete data – is the review going to be undertaken on the 'once randomised

always analysed' basis, that is, an intention-to-treat analysis?

Each of these problems has several solutions but the reviewers must choose the one that they understand and feel to be sensible. For example, dichotomous data can be summated by use of odds ratios or risk ratios, and continuous data by weighted mean difference. Different scales can be summated by using the effect size statistic (Wolf, 1986) or not summated at all. Managing lost data is difficult. Continuing with the example, the reviewers of the CPA may suggest that 'relapse' is an outcome of interest to everyone. However, in both the intervention and the standard care groups a proportion of people were 'lost to follow-up'. For dichotomous data, those who dropped out before the end of the study may be assumed to have a certain outcome (for example 30% or 50% or 100% can be assumed to have relapsed). The sensitivity of the final result to the inclusion of this assumption should be tested (sensitivity analysis). With continuous data management of 'lost data' is even more problematic. The reviewers may have to suggest a cut-off point at which they feel that the validity of the continuous data to be too much under threat (for example, 30% incomplete).

Conclusion

Systematic reviews of trials are hard work. They provide an opportunity to look at data in a quantitative as well as a qualitative way. This approach not only has an intrinsic appeal to the scientist in us, but also is supported by evidence that the effort is of value. A systematic review of case management already exists (Marshall *et al*, 1997). It shows that low-intensity case management slightly decreases the chance of a seriously mentally ill person being 'lost' to community follow-up but almost doubles the number of hospital admissions.

Systematic reviews are rewarding, enjoyable, highly regarded (Higher Education Funding Council for England, 1996), original research, that will help clarify what we do and do not know about the care of mentally ill people. The Cochrane Collaboration is still in its infancy and much work is needed before mental health has met Archie Cochrane's challenge and has systematically summarised the best available evidence of what helps or harms those with psychiatric, psychological or behavioural problems. Until that time, clinicians will continue to have problems answering the most reasonable and straightforward of questions posed by carers or those with mental illnesses. For example, if you were

offering someone haloperidol for schizophrenia and they asked, "why should I take this, what are my chances, my betting odds, of being well in six months and of getting repetitive, disfiguring facial movements in that time?", what would you say?

Acknowledgements

Thank you to Dr Iain Chalmers for advice and support with this paper and Ms Leanne Roberts for typing the manuscript. The Mental Health Foundation helped support the early work of the Cochrane Schizophrenia Group. Dr Adams is currently funded partly by the MRC, UK, and partly by the Anglia and Oxford Regional Health Authority. Dr Soares is funded by CAPES, Ministry of Education, Brazil.

Appendix: useful addresses

Cochrane Library Update Software, PO Box 696, Oxford, OX2 7YX, UK

Tel: +44 (0)1865 513902 Fax: +44 (0)1865 516918
e-mail: update@cochrane.co.uk
WWW: <http://update.cochrane.co.uk/info/>

Cochrane Schizophrenia Group Ms Leanne Roberts, Cochrane Schizophrenia Group, Institute of Health Sciences, Old Road, Headington, Oxford, OX3 7LF, UK

Tel: +44 (0)1865 226777 Fax: +44 (0)1865 226752
e-mail: csgadmin@psych.ox.ac.uk

Depression and Neurosis Group Dr Mark Oakley-Browne, Department of Psychological Medicine, Christchurch School of Medicine, PO Box 4345, Christchurch, New Zealand

Tel: +64 3 372 0404 Fax: +64 3 372 0407
e-mail: mark@chmeds.ac.nz

Psychiatry and Behavioural Science University of Auckland, Private Bag 92019, Auckland, New Zealand

Tel: +64 9 373 7599 Fax: +64 9 373 7493
e-mail: m.oakley-browne@auckland.ac.nz

UK Cochrane Centre (mental health issues) Dr Clive E Adams, UK Cochrane Centre, Summertown Pavilion, Middle Way, Oxford, OX2 7LG, UK

Tel: +44 (0)1865 516300 Fax: +44 (0)1865 516311
e-mail: ceadams@cochrane.co.uk

References

- Adams, C. E., Lefebvre, C. & Chalmers, I. (1992) Difficulty with MEDLINE searches for randomised controlled trials. *Lancet*, **340**, 915–916.
- , Power, A., Frederick, K., *et al* (1994) An investigation of the adequacy of Medline searches for randomized controlled trials (RCTs) of the effects of mental health care. *Psychological Medicine*, **24**, 741–748.
- Antman, E. M., Lau, J., Kupelnick, B., *et al* (1992) A comparison of results of meta-analysis of randomized controlled trials and recommendations of clinical experts. Treatment for myocardial infarction. *Journal of the American Medical Association*, **268**, 240–248.
- Chalmers, I. (1989) Evaluating the effects of care during pregnancy and childbirth. In *Effective Care in Pregnancy and Childbirth* (eds I. Chalmers, M. Enkin & M. J. N. C. Keirse), pp. 3–38. Oxford: Oxford University Press.
- , Hetherington, J., Newdick, M., *et al* (1986) The Oxford Database of Perinatal Trials: developing a register of published reports of controlled trials. *Controlled Clinical Trials*, **7**, 306–324.
- , Enkin, M., Keirse, M. J. N. C. (eds) (1989) *Effective Care in Pregnancy and Childbirth*. Oxford: Oxford University Press.
- , Dickersin, K. & Chalmers, T. C. (1992) Getting to grips with Archie Cochrane's agenda. *British Medical Journal*, **305**, 786–788.
- Chalmers, T. C., Celano, P., Sacks, H. S., *et al* (1983) Bias in treatment assignment in controlled clinical trials. *New England Journal of Medicine*, **309**, 1358–1361.
- Cochrane, A. L. (1979) 1931–1971: A critical review, with particular reference to the medical profession. In *Medicines for the Year 2000* (ed. G. Teeling-Smith), pp. 1–11. London: Office of Health Economics.
- Cochrane Collaboration Schizophrenia Review Group (1994) Schizophrenia and the Cochrane Collaboration. *Schizophrenia Research*, **13**, 185–188.
- Department of Health (1991) *The Care Programme Approach for People with a Mental Illness Referred to the Special Psychiatric Services*. London: HMSO for the Department of Health.
- Dickersin, K., Scherer, R. & Lefebvre, C. (1994) Identifying relevant studies for systematic reviews. *British Medical Journal*, **309**, 1286–1291.
- Enkin, M. W., Keirse, M. J. N. C., Renfrew, M. J., *et al* (eds) (1992) *Oxford Database of Perinatal Trials*. Oxford: Oxford Electronic Publishing/Oxford University Press.
- , —, —, *et al* (eds) (1993) *Cochrane Childbirth and Pregnancy Database*. Oxford: Update Software.
- , —, —, *et al* (1996) *Guide to Effective Care in Pregnancy*. Oxford: Oxford University Press.
- Hay, P. J. & Adams, C. E. (1997) The efficiency of searches for randomised controlled trials in the *International Journal of Eating Disorders*: a comparison of hand-searching, Embase and PsycLit. *Health Libraries Review*, in press.
- Higher Education Funding Council for England (1995) *1996 Research Assessment Exercise – RAE96 3/95*, pp. 11–24. Bristol: Northavon House.
- Jadad, A. R., Moore, A., Carrol, D., *et al* (1996) Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials*, **17**, 1–12.
- Marshall, M., Gray, A., Lockwood, A., *et al* (1996) Effectiveness of case management for people with severe mental disorders. In *Schizophrenia Module* (eds C. Adams, J.J. Mari & P. White). The Cochrane Database of Systematic Reviews (available in the Cochrane Library). London: BMJ Publishing.
- Mulrow, C. D. (1987) The medical review article: state of the science. *Annals of Internal Medicine*, **106**, 485–488.
- Naylor, D. (1995) Grey zones of clinical practices: some limits to evidence based medicine. *Lancet*, **345**, 840–842.
- Overall, J. E. & Gorham, D. R. (1962) The Brief Psychiatric Rating Scale. *Psychological Reports*, **10**, 799–812.
- Sackett, D. L. (ed.) (1994) *The Cochrane Collaboration Handbook*. Oxford: Cochrane Collaboration.
- Schulz, K. F., Chalmers, I., Grimes, D. A., *et al* (1994) Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynaecology journals. *Journal of the American Medical Association*, **272**, 125–128.
- Wolf, F. M. (1986). *Meta-Analysis: Quantitative Methods for Research Synthesis*. London: Sage.