Cochrane Schizophrenia Group

The University of Nottingham

THE COCHRANE COLLABORATION®

Setting up and maintaining a study-based register

This document aims to describe how TSCs run their study-based registers. Collaborative Review Groups have set up and run study-based registers quite differently from each other. These differences are described and explained with contributions from TSCs from different CRGs about their experiences and working practice. Each group has set up their register to suit their needs and make best use of their skills, experience and resources, while satisfying requirements for register to CENTRAL and PsiTri (for Mental Health groups).

The document named 'Study-based registers – a brief introduction' describes what studybased registers are and their benefits.

1. What do study-based TSCs do?

Adding and editing references and studies

- References from electronic searches and handsearching are checked for duplicates and (from the full text or abstract) if they are reports of RCT/CCTs they are entered into the register. Some groups code the study design (RCT/CCT) at this stage, others code the study design along with other study information at the next stage.

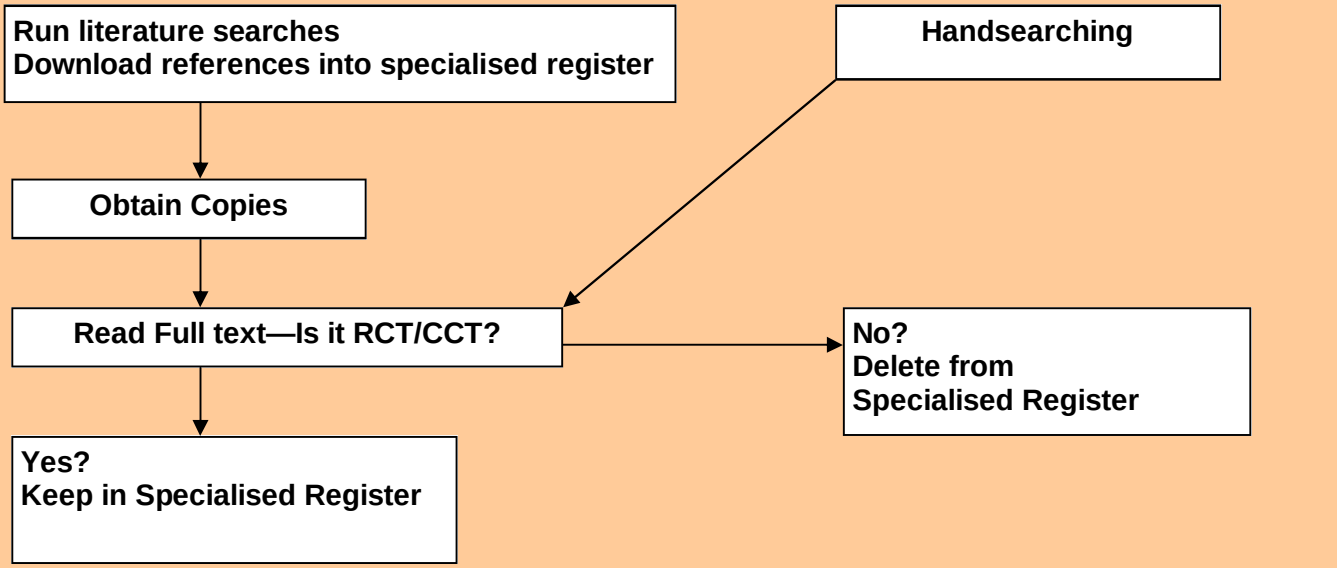
- The record is 'coded' based on the full text – the TSC picks out key study data from the paper e.g. interventions, number of participants, healthcare condition.

The next stage of inputting the coding and linking references to studies varies between TSCs using different software and different working practices.

- Either
  - Check if the reference is a report of an existing study, if so, link the reference to the study and add any supplementary study information found. If it is not clearly a report of an existing study, then a new study record is created and the coding information (like health condition, number of participants etc) is input into this study record. Once a review is completed, TSCs can use the groupings made by the authors of references into studies to update their registers further.
- Or
  - A new study record is automatically created for each new reference record that enters the register. Coding is input into the study record and minimal attempts are made to check if it is related to an existing study. Once a review has been completed, the TSC uses the groupings of references into studies in the RevMan file to identify which references to link together. For example, if four references are multiple publications of the same study, then the four corresponding study records are 'merged' into one study record, pooling the coding data. Inclusion of studies within reviews is tracked at the same time. Some well known studies are identified at the coding stage. In that case the TSC links the reference to the existing known study and adds any further study information to the record. This is the usual method for MeerKat registers

Figures 2 and 3 describe how studies are created and edited in study-based registers.

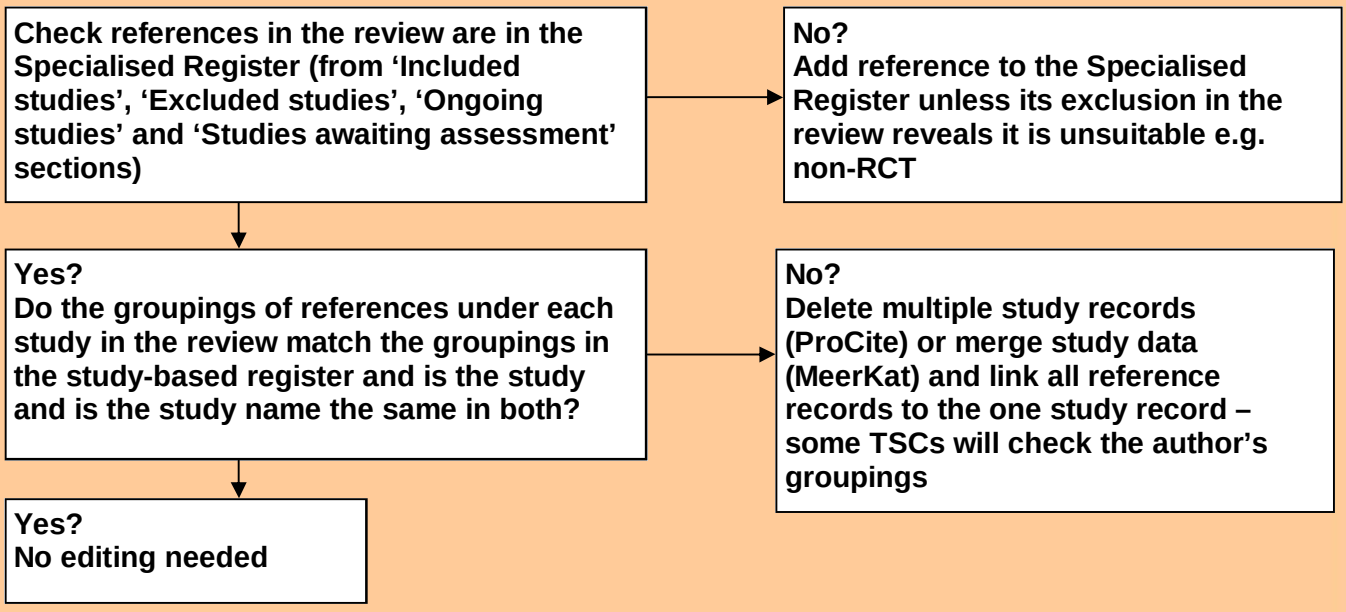
Figure 2 Adding references and studies into a study-based register



Code study data – identify interventions, health care conditions etc…
Check if it looks like an existing study
Add further study info to existing study OR Input study info in new study record

[The order of these steps varies between TSCs depending on software and preferred working practise]

Figure 3 Follow-up of references and studies identified in reviews



2. Searching a study-based register for a review

Searching study-based registers is different to reference-based registers because it must include:

- all studies that are fully-coded and linked to all multiple reports
- any studies that are coded but not yet 'merged' into the unique study that contains links to all multiple reports (in MeerKat)
- any references that are not yet coded

TSCs may need to run more than one search to ensure all relevant references and studies are found regardless of where they are in the process.

3. Setting up a study-based register

3.1 Transfer from a reference-based register to a study-based register in MeerKat

The Schizophrenia group can advise and support TSCs with this process. The amount of work and time needed can vary depending on the content and structure of your existing register. Usually it takes about 2 weeks to generate a study-based register in MeerKat from a reference-based register in bibliographic software such as ProCite. At this stage, each reference is linked to a unique study containing any coding brought from the original reference-based register. This time estimate does not include further coding required or linking references and merging of duplicate studies.

Files from the original reference-based register are formatted so reference and study records can be automatically generated in the new MeerKat register. The reference details are transferred into the reference record and any study coding in the original register is transferred into the attached study record

The transfer process can be quick and problem free if:

- Reference fields are similar to those in MeerKat
- Any study coding in the original reference records is held in separate fields to the reference coding. If the study coding is found in fields with reference coding then this first has to be separated
- The data are 'clean'. Inconsistencies in formatting e.g. Author names and initials, Journal names, Interventions become obvious. MeerKat can export references for import into RevMan, so it is important they are accurate for the reviews. Registers are 'cleaned up' for the transfer – this is a good opportunity to improve your register and get help with the process.

Once a review has been completed, the TSC can merge (update and rationalise) studies found in their register according to the groupings of references made in the review. For a more 'studified' register a retrospective look through completed reviews reveals identified studies and their associated references. The study records can be updated with the links between references made in published reviews. Any references not yet included in any reviews would then be looked at to see if they are multiple reports of studies.

3.2 Changeover from a reference-based register to a study-based register in Procite

- The principles of running a study-based ProCite register are exactly the same as for MeerKat and the results are similar, see PsiTri (http://psitri.stakes.fi/index.html). There is no difference between the records which originate from MeerKat and the ones that originate from ProCite.

- You will need to use different ProCite workforms from the ones you may be used to and you may have to move fields in order to fit in with that (this can be done by global edit and is quick and easy).

- The existing PsiTri manual has detailed instructions on how set up ProCite to go study-based and on how to go about creating study records. Within a day you will be fluent at this. There are no new skills to learn and you can run your register in much the same way as you used to.

ProCite study-based registers have been set up in two ways:

- A single ProCite register containing reference records and study records. You can work on the study-based part of your register as and when you have time. Your reference and study records are all in one database; when you search you will get 'studified' and 'unstudified' records, and coded and un-coded records so your authors will get the benefit of any study-based records right from the start.

- Two ProCite registers running side by side containing references and study records. This register is described more fully in the case studies section (5.2 CCDAN).

4. Why have different groups set up study-based registers using different software?

Bone, Joint and Muscle Trauma Group (formerly Musculoskeletal Injuries)

Reference Manager was selected as the software for the Group's study-based Specialised Register simply because it was supported by the University of Edinburgh, where the Group was based at that time.

Depression, Anxiety and Neurosis Group (CCDAN) choice for ProCite for CCDANCTR Studies

On agreeing to become involved in the EU-PSI project we ran a pilot using two inexperienced coders. We asked the coders to code 50 references each, 25 in ProCite and 25 in MeerKat with the order reversed. We asked the coders to record the time needed to code each record and the user-friendliness of each software system.

On completion of this pilot ProCite was chosen on the basis that it has more fields easily available for use. ProCite study-based registers can be easily submitted to and complied by PsiTri staff for the EU-PSI trials register.

Dementia and Cognitive Impairment Group (CDCIG) choice of ProCite over MeerKat for a study-based register

The decision to remain with ProCite for a study-based register was based on two reasons. The first was that MeerKat at the time was not supported by the Collaboration (and still is not) and the second reason was a practical one: keeping a study-based register in ProCite was easier to learn and less time-consuming. For the work that TSCs do it is adequate and gives results that are no different from those you get with MeerKat (PsiTri shows you that).

Schizophrenia Group preference for MeerKat over ProCite for a study-based register

ProCite and MeerKat are two very different pieces of software, which do very different jobs. MeerKat is specifically designed to manage study information and in our view is better at linking all the pieces of information, such as reference to study, study to review, and reference to reviewer.

MeerKat is a sophisticated Microsoft Access-based database. The flexibility of Access allows the data in the register to be used and manipulated in other programs. The benefits of MeerKat to the Schizophrenia group have been the ability to customise and develop the system with limitless fields and the development of 'sister' programmes to help TSCs work. For example the 'Full text finder' (Ferret) checks the references in a MeerKat register that a group has not yet got a hard copy of, and produces a report listing the web sites, libraries and even library shelf marks where they can be found.

Being a ‘relational’ database means it can be used as a powerful tool in tracking references and studies used by specific authors in specific reviews. There are other relationships between the data in our file, automatic date insertion and global editing to control indexing and coding terms, all of which help keep the register clean with relatively low maintenance.

Stroke Group

The Stroke Group was registered with the Collaboration in August 1993 and shortly thereafter decided to convert their existing database in Reference Manager, of just over 1300 into a study-based register. At that time no specific software existed for this purpose, and the Stroke Group decided to develop its own - RefTraK.

5. Study-based register case studies

5.1 Bone, Joint and Muscle Trauma Group (formerly Musculoskeletal Injuries)

Reference Manager

Lesley Gillespie

The Group was originally established in 1995 as a subgroup affiliated to the Musculoskeletal Group but an application in July 1998 for retrospective registration for full Group status was successful.

We have a coding system using keywords to code for method of randomisation, condition, type of intervention (surgical, conservative etc), specific comparisons, participant numbers and age range. The register has always been study-based in that keywords are used to link references that relate to the same trial. A code (keyword) is assigned for the condition and its site in the body and whether it is treatment or prevention. These are numbered in sequence. Thus the first paper relating to a trial for the treatment of intracapsular hip fractures was coded ‘treat.femur.fractp.intra001’ and any subsequent papers that related to the same trial were assigned the same code, so all references relating to that trial can be retrieved.

With this method you do have to put the coding in for each paper. Sometimes you can just copy and paste all the codes across from the original paper, however, you cannot assume that the details will be the same in each paper as some report different outcomes, and often the sample sizes will differ.

We try to group into studies as references are identified (prospectively with regards to reviews) but I also check new and updated reviews to see whether review authors have identified new references relating to existing trials (retrospectively with relation to reviews). If they have I would then add them into Reference Manager.

I would say we use this coding and grouping into studies more as a double check when reviews come in (to check that authors have grouped separate references as one study), or in order to identify areas for new reviews. We feel the final decision about which references relate to any one study lies with the authors of the reviews as one of them is the guarantor.

Many of our authors would write to the authors of trial reports and will obtain more information than is available in the published papers. To code on the basis of the papers alone could then be wrong. It can be quite difficult to determine whether separate papers relate to the same trial or not: if two papers with the same methods have a different sample size one might be an interim analysis or it may be a separate pilot for the larger study, i.e. a separate study. Sometimes it is necessary to correspond with authors to determine whether separate references relate to the same study i.e. resource intensive and perhaps best left to the review authors who are corresponding anyway to obtain further details of the randomisation method etc.

5.2 Depression, Anxiety and Neurosis Group (CCDAN)

ProCite

Hugh McGuire

CCDAN runs two registers in ProCite.

CCDANCTR-References
This register currently contains more than 16,011 references referring to completed or ongoing trials.

CCDANCTR-Studies
We have coded 12,004 references into 9,095 studies.

Details of trial coding
The following fields have been coded using the EU-PSI coding manual. (Eu-Psi Coding Manual) Study Name, Current Status of Trial, Study Design, Blinding, Intervention, Adjunctive Intervention, Number of Participants, Unit of Allocation, Link to References, Link to Cochrane Reviews, Health Condition, Co-morbid Health Condition, Country of Origin, Treatment Setting, Main Diagnostic Criteria, Age, Duration of Trial, Anticipated End Date, Washout/Run-in, Unchanged Study Sample Size, Information on Dropouts/Withdrawals, Outcomes, Funding Source/Sponsorship, Side Effects, Methods of Treatment Allocation, Economic Evaluation.

Cochrane Schizophrenia Group

The University of Nottingham

Coding process

1/ Trial reports identified from various sources. Electronic downloads held in a ProCite database, de-duplicated and screened for relevance to CCDAN.

2/ If relevant the record is entered into CCDANCTR-References.

3/ Hardcopies are then obtained and checked again for relevance to CCDAN. If not relevant then record is taken out from CCDANCTR-References and submitted to the Cochrane Library as a handsearch result and the hardcopy article is forwarded to the relevant Cochrane Group.

4/ The reference is copied to CCDANCTR-Studies where a check is made to see if the reported trial already exists. If so the reference is added and a link added to the reference record.

5/ If the report is to a new trial, the trial information is entered and the reference is added and a link added to the reference record. Individual references may be linked to more than one trial.

Searching process

When a new title is registered or protocol received, the TSC and author decide on the best search strategy to use, CCDANCTR-Studies alone, CCDANCTR-References alone or both together.

Searches in CCDANCTR-Studies are based on the “Criteria for considering studies for this review” section with as many patient, intervention, comparison and outcome terms as needed.

Searches in CCDANCTR-References are based more on keywords downloaded with the reference and free-text terms. CCDAN can also check the un-coded references in CCDANCTR-References as we can search abstracts (which is not possible in CCDANCTRStudies)

Typically a CCDAN search will include both. First the CCDANCTR-Studies search will be carried out and the only the “un-coded” references CCDANCTR-References will be searched. These “un-coded” references can then either be forwarded to the author and prioritised for coding or – time permitting – be coded immediately.

Finally when the review is submitted for editorial scrutiny the TSC can quickly check if studies identified by the search strategy have been included or excluded.

PDF created

5.3 Schizophrenia Group

MeerKat

Mark Fenton and Judy Wright

Transferring from ProCite to MeerKat

The Schizophrenia Group has always had a study-based register of sorts since its inception in 1994, (it was a flat file, with inconsistent attempts to link to studies), which until 2001 was in ProCite.

Back in 2001 we learnt the hard way, transferred fifty records at a time, cutting and pasting the study information into our MeerKat register over a three to four month period. We can now transfer a whole register, with all its coding, over about two weeks and have helped transfer many other groups registers from a variety of bibliographic software into MeerKat. We can offer advice and support where possible for those wanting to move their registers to MeerKat

Using MeerKat

We continue to use ProCite to import from MEDLINE etc, duplicate check, and format fields for a clean export to MeerKat.

We currently have over 8173 references reporting on 6452 studies, the actual number of studies is less than this and will fall as duplicate studies are identified and merged.

The fields coded using PsiTri manual are: Study Name, Status of Study, Study Design, Interventions, Healthcare Conditions, Outcomes, Participants (age and sex), Number Randomised, Country of Origin (of study), Link to references, Link to Cochrane Reviews.

Searches are run on both the reference 'layer' of the register and the 'study' layer to ensure all relevant studies and references are found. We track which references and studied are sent to authors, so update searches can automatically disregard items already seen by reviewers.

The Schizophrenia register is fully study-based in that it codes and creates studies for all references in the register. However, there is always a backlog of coding and processing studies identified in reviews and merging duplicate studies. We are considering developing a predictive test, to identify reports that are likely to be from the same study. This would help identify studies that the reviews have not yet included and identify duplicate studies.

A report of this register has been published:
Thornley, B. & Adams, C. (1998) Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ*, 317, 1181 -1184.

(http://bmj.bmjjournals.com/cgi/content/full/317/7167/1181?maxtoshow=&HITS=10&hits=10&RESULTFORMAT=&author1=thornley&andorexactfulltext=and&searchid=1113388938440_6599&stored_search=&FIRSTINDEX=0&sortspec=relevance&resourcetype=1)

5.4 Stroke Group - Specialised Register Management Software

RefTraK

Brenda Thomas

History

The Stroke Group was registered with the Collaboration in August 1993 and very shortly thereafter we decided to convert our existing Reference Manager database of just over 1300 references to trials, into a study-based register. Based on experience from the Antiplatelet Trialists' Collaboration it was clear that unless all publications from individual trials were linked, the possibility existed of including data from the same trial more than once in a systematic review. We also recognised that with a study-based register we would be able to reduce the workload for authors and more easily monitor the authors' response to trials sent. At that time no specific software existed for this purpose, and the Stroke Group decided to develop its own - RefTraK. This decision has had major implications for the development of the Group.

Functions of RefTraK

- A relational database containing review information and bibliographic and trial details, where the TRIAL is the unit of data
- Uses a sophisticated hierarchical coding system for interventions
- Can link multiple reports to single trial
- Helps track trial progress (and duplicate publication)
- Can identify interventions where reviews are needed
- Provides authors with regular lists of new and relevant trials
- Identifies out-of date reviews: lists new trials since last review update

Overview of process

(a) Trial reports identified from various sources.
Electronic downloads held in Reference Manager databases and screened for relevance. Hard copies obtained of all potentially relevant trial reports and further assessed for inclusion in trials register. Copies obtained for all relevant trials identified by handsearching.

(b). Each trial report checked against RefTraK.
If trial already exists, reference is added and linked. (The paper copy archive is useful here to check references are from the same trial). At this point trial details may be updated. If report is to a new trial, the trial information is entered and the reference is added and linked. Individual references may be linked to more than one trial.

(c) When a new review is registered the details are entered into RefTraK and linked to specific interventions which are saved with the review. Regular searches are run for each review based on trial coding elements (eg intervention, disease stage, control etc). Authors receive details of new trials and associated references since last search plus details of new references to trials already sent.

(d) Authors complete Author Response Form assessing relevance of each trial sent. It is therefore possible to check that reviews are being kept up to date.

(e) Once review is submitted new trials and new references to existing trials found by authors from other sources are added to the trials register.

Implications for Stroke Group

FUNDING REQUIREMENTS

(a) Computer Programmer required for maintenance and further development of RefTraK
(b) Full-time TSC required to combine trial coding with trial identification
(c) Additional cost of obtaining hard copies of trial reports (currently over 9000 trial reports in register)

ADDITIONAL COMMENTS

The development of RefTraK is one of a number of ways of managing a trials register. The Stroke Group has committed resources to its development and maintenance and it works well for us. We are currently developing a new faster more efficient version. Other groups will have different resources, priorities and mechanisms to support authors.

The Trials Register is now incorporated into EffectiveStrokeCare.org, a fully indexed, searchable, web-enabled database of evidence for stroke care management (www.EffectiveStrokeCare.org)

5.5 Cochrane Dementia and Cognitive Improvement Group

ProCite register

Dymphna Hermans

History

We took part in an EU project to create a Mental Health Library which ran from 2000 to 2003/4. Part of this project involved the creation of a study-based register of trials in the field of mental health (PsiTri).

We obtained funding to employ a full time coder for a period of 18 month to work his/her way through the 3000 odd references in our SR. We had to pay for obtaining hard copy ourselves and decided not to attend the Colloquium in Capetown in order to be able to obtain them. We had calculated the coder would be able to code and link on average 10 references a day, 50 a week, 200 a month, 2400 a year, 3000 in 18 months. We completed the coding on schedule. Trials in dementia tend to be relatively straightforward in their design. Trials in other areas might take longer to code and link.

Details of trial coding

Trial records have the following fields: status of trial, study design, blinding, intervention, absolute number of participants, unit of allocation, health condition, co-morbid health condition, country of origin, main diagnostic criteria, ISRCTN, outcomes.
For searching purposes they have proved to be sufficient. We have started to code duration of trial, because that is very helpful for linking references to studies.

Overview of process

Records from electronic downloads and handsearching are downloaded in our ProCite SR and screened for relevance: on the basis of the abstract we code as RCT or CCT (if in doubt, we include at this stage). Hard and e-copies are obtained of all new trial reports.

Then as and when we have time, we 'studyfy: we search for references to the same study before we start coding. If we find it is a duplicate reference, we edit the study record if necessary; if it is a unique reference we create a trial record and code the details of the trial.

When the first draft of a review comes in, the TSC checks the included and excluded studies against the SR and against the search reports sent to the reviewer and solves any discrepancies (sometimes after discussion with the authors; sometimes we are right; sometimes they are) before the review is seen by anybody else at the editorial base. New trials and new references to existing trials found by authors themselves are added to the trials register if they were not already there.

It is essential that the study names used in the review are the same as in the SR as in PsiTri studies are linked to reviews by the Helsinki editorial base and they can only do that if the study names are identical (it can be quite a lot of work to keep these synchronised).

Both study and reference records have a field for update codes (it can have as many codes as you like but separator must be used). All records that have been sent to authors in the past have this code (e.g. Donepezil for AD is DON). The last batch of records (studies and references) sent to an author have the code in the format DON-sent. When a new search is carried out the DON-sent code is removed but the DON code retained. The new results get the DON-sent code.

So:
DON: authors have seen this record in the past and dealt with it
DON-sent: the last batch of records sent and being dealt with
No code: records have not been seen by the authors yet.

Funding requirements
Fundraising will be essential if we are to continue to keep our study based register up-to-date now that EU funding has ended. The cost of obtaining hard copy is an additional burden although the proliferation of free e-journals is helpful.

Additional comments
The Dementia group has committed resources to the development and maintenance the ProCite study based register and it works well for us. Other groups will have different resources, priorities and mechanisms to support authors.

Document prepared by:

Judy Wright,
Trials Search Co-ordinator,
Cochrane Schizophrenia Group
School of Community Health Sciences
Duncan Macmillan House
Portchester Road
Nottingham
NG3 6AA
United Kingdom

Tel: +44 (0) 113 343 1897
Fax: +44 (0) 113 343 2723

Contributions by:

Kathie Clark,	Canadian Cochrane Center
Mark Fenton,	Schizophrenia Group
Lesley Gillespie,	Bone, Joint and Muscle Trauma Group
Dymphna Hermans,	Dementia and Cognitive Impairment Group
Carol Lefebvre,	UK Cochrane Centre
Hugh McGuire,	Depression, Anxiety and Neurosis Group
Victoria Pennick,	Back Group
Brenda Thomas	Stroke Group
Matthew Parker	Schizophrenia Group