



Study-based registers – a brief introduction

The extent to which individual registers are ‘study-based’ and the methods for achieving this are very variable. This document intends to offer the basics of study-based registers. The document named ‘Setting up and maintaining a study-based register’ reveals in more detail the work of study-based TSCs and management of their study-based registers in different review groups.

What are study-based registers?

Study-based registers contain the same information as reference-based registers but link together references that report on the same study.

Study-based registers often (but not always) store additional information that describes the study that the reference/s report. This additional ‘study information’ varies depending on the needs of the review group.

Study information is also referred to as ‘coding’ and can be held in either:

- user defined fields of reference records in a reference-based register OR
 - unique study records which link to the related references in a study-based register.
- ‘Coding’ involves reading and picking out key study information, usually from the full article, and entering these data into the specialised register.

A reference-based register is a flat file of bibliographic data organised into fields and references. Some study-based registers are ‘relational’ databases. Fields can be related to each other, allowing references, studies, reviews and authors to be inter-linked, making it a powerful tool for managing references and studies.

Are all records ‘studified’ in a study-based register?

Study-based registers (like reference-based registers) are ‘work in progress’ as TSCs try to keep their register as up-to-date as possible. They aim to contain a set of references that are all neatly linked into distinct studies (i.e. fully studified). In reality the registers are constantly changing as new references are found and references can be retrospectively grouped into studies. The scope of the group and their available resources affects the backlog of sourcing, coding and linking references into distinct studies. Work can be prioritised by first dealing with references and studies that have been or plan to be used in reviews.

Who uses them?

All TSCs were asked via the TSC mailing list to state if their register is study based (i.e. with links between references of the same study). From 37 responses, 12 said yes and a few stated they have some cross-referencing of references to studies but it is not being done systematically. All study-based groups code study data, though for a few this is just identifying if the study is an RCT or CCT, not identifying data such as interventions, participants, healthcare conditions and outcomes. Study-based registers have been developed using MeerKat, ProCite, Reference Manager and RefTrak software.

What do they look like?

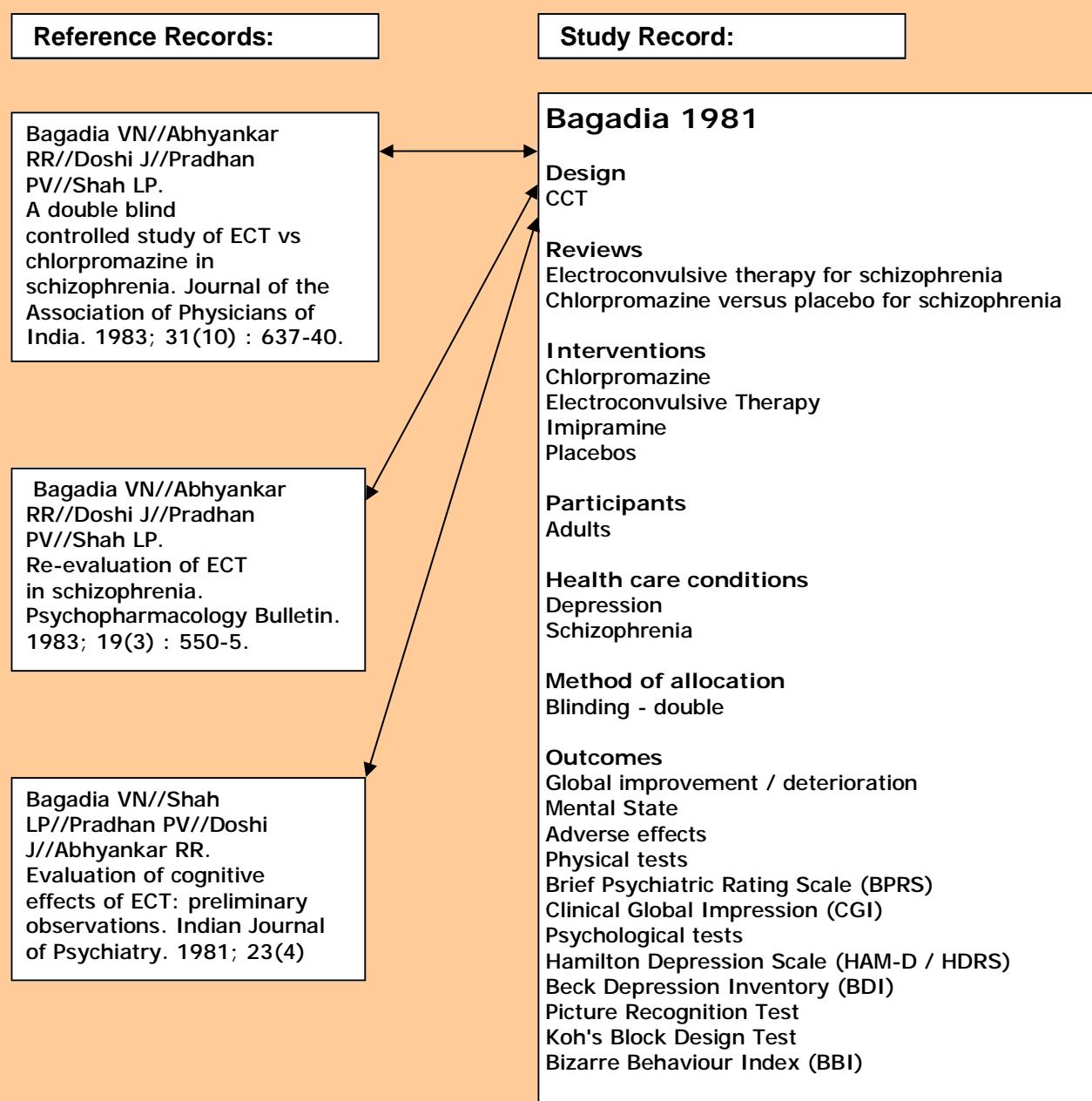
PsiTri (Mental Health Library) is a working example of a study-based register. It is freely available on the Internet and was created as part of the [EU-PSI project](#) involving Mental Health CRG registers. As part of the project the CRGs involved helped produce a coding manual. The manual aims to ensure consistency in coding study fields and the format of references across the CRGs.

<http://psitri.stakes.fi/index.html>

To visualise a register containing both reference and study records it may be helpful to think of two 'layers' of records in the database. One layer is the reference records and the other is the study records. Each study record contains information from, and links back to the references that report it.

Figure 1 displays the typical contents and structure of the study-based register, for one study which has been reported three times.

Figure 1. Structure and contents of Bagadia 1981 study with 3 reference records attached



2. Benefits of study-based registers

2.1 Benefits to authors

The reduction of authors 're-inventing the wheel' when grouping references into studies. When an author is looking at references that have not been related to a study, they have to start from scratch, identifying unique studies with multiple publications. Search results from study-based registers show the author the groupings of references into studies already made by other authors and (sometimes) TSCs.

Example

The Schizophrenia Group has found over 120 references that report a unique study. The study compares a number of different drugs and has been included in several reviews. The authors initially spent considerable time identifying that all these references related to the same study. They have been grouped together and recorded in the study database. Future authors, who retrieve these references in their register search results, will see that they have been identified as one study. They need to decide whether to include the study and look for further reports on the study. They do not have to look at all 120 references that have been previously identified and grouped, to determine whether they relate to the same study.

2.2 Benefits to TSCs

- Quicker and more efficient coding. For TSCs who code study characteristics, the study data is coded once for each study. Multiple reports of a study are only scanned to see if any further study data can be added to the study record. In a referencebased register the study coding is repeated for each reference that reports a particular study.
- Authors value the study search output and TSCs for managing it, as it can reduce their workload. The value of the study search output improves over time as more references are grouped into unique studies. The more 'studied' the register, the greater the value of these searches in reducing duplication of authors' effort.
- The number of studies in a given area can be accurately assessed, rather than the number of references.
- More efficient removal of non-RCT/CCT references. Once a study has been identified as not an RCT/CCT, then the study record and all related reference records are removed from the SR together. The linking of multiple reports to the study makes it very quick and easy to identify which references to remove.
- Planned and ongoing trials can be more easily tracked to full publication.

2.3 Benefits to The Cochrane Collaboration:

Cleaner Specialised Register

- Coding becomes a more efficient and less time-consuming task. The TSC has more time to code the whole of their register. This in turn will help produce a cleaner register as TSCs have chance to look at more full copies to determine RCT/CCT status when coding.
- Looking through the Excluded Studies tables in the completed Review shows why the author excluded the study. If it is a non-RCT/CCT you could remove it from your register, having re-checked from the full paper that the review author is correct. If it is not the CRG's healthcare speciality you should pass it on to a more appropriate CRG, or resubmit to CENTRAL as a Handsearch result, to ensure it is not lost from CENTRAL. Either way, it is quicker to remove the study along with all multiple reports from a study-based register, rather than search and remove the reports one-by-one from a reference-based register.

Consistency with RevMan. The organisation and linking of references to studies is in line with the RevMan organisation of references by study.

2.4 The pros and cons of coding (applicable to study-based and referencebased registers)

Pros

- A search of a coded register can select trials/references/studies more accurately meeting the inclusion criteria for a review. For example, a search for records with 'Weight change' in the outcome field would produce less spurious results than searching for 'weight' in title, abstract and index fields. Similarly a search within the coded intervention field for 'diet' terms would produce less irrelevant hits than a search for 'diet' anywhere in the bibliographic record.
- Develop a cleaner register (and therefore cleaner search results) from identifying RCT/CCTs and problem references/studies highlighted in the review.
- Improve subject knowledge and ability to identify RCT/CCTs by reading and picking out key study data when coding articles.
- The value of different sources for RCTs can become easier to see when coding sources such as conference proceedings.

Cons

- The coding process is time consuming, leaving less time for trial identification.
- Time and money costs to obtain hard copy – though this is decreasing as free access to electronic journals is improving.
- Physical space to house hard copies – though this is less of a problem if you are happy storing pdfs and you have the copyright permission to do so.

Document prepared by:

Judy Wright, Trials Search Co-ordinator, Cochrane Schizophrenia Group

Contributions by: Kathie Clark, Mark Fenton, Lesley Gillespie, Dymphna Hermans, Carol Lefebvre, Hugh McGuire, Victoria Pennick and Brenda Thomas