The University of Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

Shokraneh, Farhad and Adams, Clive Elliott (2017) Study-based registers of randomized controlled trials: starting a systematic review with data extraction or meta-analysis. Bioimpacts, 7 (4). pp. 209-217. ISSN 2228-5652
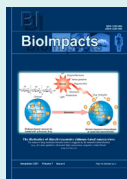
Study-based registers of randomized controlled trials: Starting a systematic review with data extraction or meta-analysis

Farhad Shokraneh[1]*, Clive Elliott Adams[1]

[1] Cochrane Schizophrenia Group, the Institute of Mental Health, a partnership between the University of Nottingham and Nottinghamshire Healthcare NHS Trust, Nottingham, United Kingdom

Abstract

Introduction: Despite years of use of study-based registers for storing reports of randomized controlled trials (RCTs), the methodology used in developing such registers/databases has not been documented. Such registers are integral to the process of scientific reviewing. We document and discuss

methodological aspects of the development and use of study-based registers. Although the content is focused on the study-based register of randomized/controlled clinical trials, this work applies to developers of databases of all sorts of studies related to the human, animals, cells, genes, and molecules.

Methods: We describe necessity, rationale, and steps for the development, utilization and maintenance of study-based registers as well as the challenges and gains for the organizations supporting systematic reviews of the published and unpublished literature.

Conclusion: The ultimate goal of having a study-based register is to facilitate efficient production of systematic reviews providing rapid, yet accurate, evidence for the decision-makers. We argue that moving towards study-based registers is an inevitable welcome direction and that infrastructures are ready for such movement.
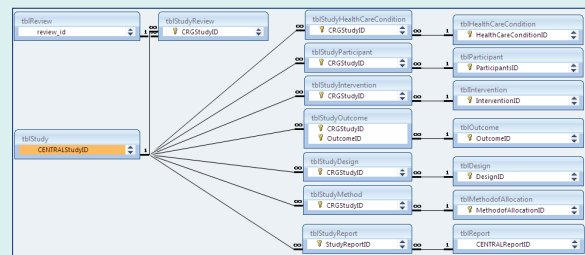
Introduction

Emergence of specialized organizations creates needs for the specific information for decision-making. Developing and maintaining specialized registers or databases is a valuable tool for achieving organizational goals by providing information to support such decision-making.[1,2] Specialized bibliographic registers save time by allowing specific and sensitive searches for records of documents of high relevance to the research question.[3]

The Cochrane Collaboration (1993 – present) is an international non-for-profit network of, amongst others, healthcare researchers, practitioners, and patients. This Collaboration works to summarize and synthesize available evidence to support informed decision-making in the healthcare sectors. The Collaboration's main activity is producing high-quality systematic reviews of randomized controlled trials (RCTs).[4] However, over the

last two decades after its foundation, there are still only 6230 reviews published on the Cochrane Library (June 2017), and most are in need of update. Despite enormous efforts of thousands of researchers, there remain hundreds of thousands of clinically valuable randomized trials not summarized within reviews. Best evidence to support the health care decision-making is being wasted. It is not an impossible task to provide good coverage of all clinically useful and best evidence of the effects of health care in any one sub-specialty but it is a large and difficult task. Since a systematic review requires a long process including literature searching, de-duplication of search results, screening the search results, obtaining the full text of the reports, putting the publications from the same study together (studification), data extraction and meta-analysis, it has become clear that study-based registers of randomized trials are integral to this task shortening

*Corresponding author: Farhad Shokraneh, Email: Farhad.Shokraneh@nottingham.ac.uk

the process for the reviewers to start a review with data extraction or meta-analysis.[5,6]

From the very start the Cochrane Collaboration created a specialized register of relevant literature – now called the Cochrane Central Register of Controlled Trials (CENTRAL) – and has disseminated this in Cochrane Library.[7-9] CENTRAL is now the largest bibliographic reference-based database of reports of randomized trials.[10,11] It is an amalgamation of the individual Cochrane groups' registers which, in turn, are developed from biomedical bibliographic databases such as EMBASE, MEDLINE and other sources such as conference proceedings.[12] In such databases, the publications and the reports from the same study are not linked together and putting the salami of the study back together required massive efforts.

Slow production of update systematic reviews because of long process (6230 reviews in 20 years) and salami publication of trials were two main problems leading in the development of study-based registers. In recognition of the value of study-based registers for more efficient review production the methodology for development, utilization and maintenance of such register have been documented in this methodology paper.

Objectives

- To search the literature for relevant papers on study-based registers.
- To report the rationale methods of development and challenges of study-based registers for which relevant documentation seems remarkably sparse.
- To share more than two decades of practical experience in the creation and maintenance of study-based registers of biomedical literature.

Search methods

We ran a search to find relevant literature and to ensure the novelty of the current paper. Since the terminology in this topic is standard, we used the following search strategy on MEDLINE (1946 to Search Date) and EMBASE (1974 to 2017 Week 34) via Ovid SP and updated this search on August 29, 2017:

Search Strategy: ("Study-Based" adj (Register* or Database*)).ti,ab.

We also searched all conference abstracts presented in Cochrane meetings. We did not identify any full published paper on the study-based registers but did identify many conference abstracts. We then contacted Cochrane Information Specialists and followed web searching to find all possibly relevant documents for this overview.

Specialized registers
1. Reference-based registers

Within a reference-based register (database) each record usually behaves as a separate independent entity. There is one record for each 'reference' (also referred as 'citation', 'publication' or 'report'). However, a study may have several reports or publications and hence, several records in the register.[13] For instance, the researchers may conduct one study but report it in several references such as conference abstracts, a dissertation, a poster, journal papers, an online trial registry record in "ClinicalTrials.Gov" and so on. Although the references are different in the format they all present some or all data from a single study and each is listed separately within the reference-based register. A reference-based register is the least resource-intensive and simplest register to assemble.[13]

A reference record in a reference-based register may consist of bibliographic information, abstract and indexing fields. This record may also contain a link to the full report of the reference or be the full report in itself (i.e. conference abstract).[14,15] In such a register, the reference record may sometimes contain another link to other references of the same study,[16] or a study name or unique identifier, such as online trial registry number.[17] Such reference records are the backbone of all the major bibliographic biomedical databases such as MEDLINE and communication between such databases and reference management packages is usually easy. This type of register may represent the totality of publishing activity, but, because of the 'one-to-many' issue of so-called 'salami' publication of one study, the register is not an accurate representation of the total research activity.

2. Study-based registers

Conducting one study, the researchers usually published their data in more than one paper.[18] It means that one study might have several reports. To represent the whole data, we should find, use and cite the whole study rather than one paper or one reference of the study. For those working in systematic reviewing – summarizing research activity rather than publishing activity – there is the necessity to work at the study level instead of the reference level. Every time a systematic review is conducted a study-based register of some sort is created. Recognizing the wastage involved in disassembling these at the end of each review, some Cochrane groups now maintain study-based registers in which all references of one study are linked to a single meta-record called the 'study record' (also referred to as the 'trial record').

The idea of linking related records to one meta-record is not a new concept in information science[19] but it is the cornerstone of modern relational databases. At the most basic level, a study-based register consists of study records each of which links to all available references of each study.[20] Further value may be added to the study record by having it contain structured study-specific meta-data such as data about who were the participants, what are healthcare **P**roblems, what are **I**nterventions and **C**omparator, and what **O**utcomes have been measured-so-called PICO framework.[21,22] These metadata make it possible to run a very specific search for the particular

intervention relevant to a certain healthcare condition and retrieve all the related trials – and no more – at the click of a button. All the data and their associated meta-data are stored in data tables (Figure 1). One meta-record may also contain information about all relevant references. Even though a reference may have contained similar PICO meta-data there is no implicit guarantee that data in the one reference record accurately represents all activity within the study. A single reference often reports a subset of activity within the study.

The study record can also contain accurate controlled vocabulary[26] describing the comparisons within the study (e.g. drug X versus drug Y for condition Z) with the potential for direct linking to the end review(s).[27] Finally, the study record may also contain all tabulated data extracted from the full-text reports or the full primary dataset (Table 1).[28]

Rationale

The 'unit of currency' of systematic reviews of healthcare is the study – not a single reference or a single published paper. The undertaking of a systematic review necessitates the creation of some sort of study-based list or register whether it be by the reviewers or Information Specialist. If pre-prepared by the Information Specialist the final product can be more accurate, save the time of other researchers, and has the potential to be re-used to avoid duplication of effort in the future. Certainly, if the single reference is considered the main unit for systematic reviews, erroneous double or even triple counting of data from various references of a single study is a real danger (Box 1). The use of study-based registers decreases this danger – although it does not make it impossible.[29]

1. Reference-based register or study-based register?

Some research groups prefer to keep their trials register at a reference level. However, many in systematic reviewing research have made (what we feel to be) the inevitable move to becoming study-based. In reality, one has to underlie the other. The study-based register should sit over a clean repository of references.[33-35]

Moving to a study-based register depends on the research teams' ethos and policies, planning and available resources. Holding a reference-based register will save search and de-duplication time. However, a study-based register allows Information Specialists not only to save that searching and de-duplicating time but also can benefit authors of reviews by providing study records with multiple references already organized at study level. Study-based registers also have the capacity to save author-time by supplying records in which qualitative and numerical data are extracted and referenced.

2. Comparing functionality

Because of the high relevance of screened references to the methodology and healthcare condition, having a specialized reference-based register does increase the sensitivity and specificity of searching for reviews.[3] However, the processes of searching the register, obtaining full-text reports of references, screening search results, checking each reference based on inclusion criteria, assigning references to relevant studies, concatenating data from the same study, and coding information/data to link references to study[36] take time – mostly reviewers' time. A study-based register has, to a greater or lesser extent, already undertaken at least some of those tasks moving the expense down the line to Information Specialists who are more skilled in such types of information management.

Making the move

The Cochrane Dementia and Cognitive Impairment Group spent 3 years converting their specialized register
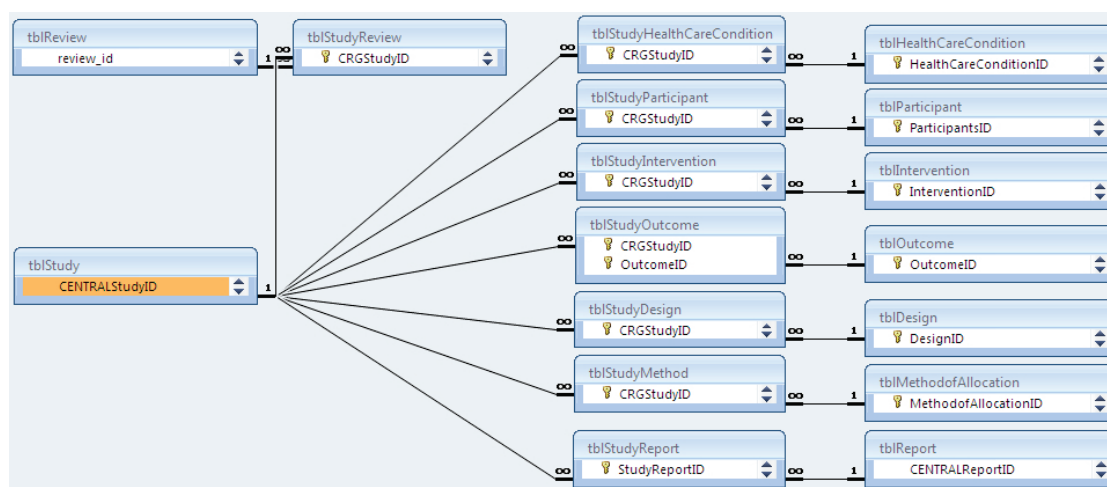
Figure 1. Example of the relationships within a real world study-based register. This figure shows 17 tables out of 34 tables (tbl) within MeerKat, a Microsoft Access-based relational database.[23-25] Bibliographic information of each report (in tblReport) is linked to its study (in tblStudy). tblStudy is linked to PICO metadata (in tblHealthCareCondition, tblIntervention, tblOutcome) and to relevant systematic review/s (in tblReview).

Table 1. A study record containing the metadata, data extracted from a trial report, and the location of the data within the report

Metadata	Data from report(s)	Location
Study Name*	Jahanian 2014	
Reference(s)	1. [Ref. ID 19855] Jahanian AA, Rezaei O, Fadai F, Yaraghchi A. The Effectiveness of Rivastigmine in Reducing Tardive Dyskinesia Symptoms in Patients with Schizophrenia. Iranian Journal of Psychiatry and Clinical Psychology 2014; 20(1): 29-34. 2. [Ref. ID 58435] IRCT2012092910964N1. The effectiveness of Rivastigmine on reducing the symptoms of Tardive dyskinesia in patients with Schizophrenia. Available from [Accessed 29 August 2017]: http://www.irct.ir/searchresult.php?id=10964&number=1	
Characteristics		Location in PDF**
Methods	Allocation: "randomly assigned" no details reported.	19855PG30C1P3L7
	Blindness: "double blind" no details reported.	19855PG30C1P3L3
	Design: not reported.	
	Duration: "eight weeks".	19855PG31C1P2L4
	Setting: "Razi Psychiatric Center, Tehran, Iran".	19855PG30C1P3L5
Participants	Diagnosis: Patients with schizophrenia and tardive dyskinesia (TD) based on DSM-IV-TR diagnosed by psychiatrist.	19855PG30C1P3L12-13
	N=40.	19855PG30C1P3L5
	Age: range 18-65 years.	19855PG30C1P3L17
	Sex: not reported.	
Interventions	1. Rivastigmine: dose: 1.5 mg twice daily. N=20.	19855PG30C1P3L7-8
	2. Placebo: no details reported. N=20.	19855PG30C1P3L10
Outcomes	TD symptoms: no improvement (AIMS).	19855PG31C1P2L5
Notes	Sponsorship source: "no financial support".	19855PG33C2P3L1-2
Risk of Bias		
Bias	Support Statement from Report	
Random sequence generation	"Randomly". No details.	19855PG30C1P3L7
Allocation concealment	Not reported.	
Blinding of participants and personnel	"Double blind". No details.	19855PG30C1P3L3
Blinding of outcome assessment	"Double blind". No details.	19855PG30C1P3L3
Incomplete outcome data	Not reported.	
Selective reporting	None. Registered protocol is available (IRCT2012092910964N1).	19855PG30C1P3L1
Other biases	None known.	
Outcome		

	Rivastigmine		Placebo		
	Mean	SD	Mean	SD	
AIMS after Intervention	12.5	7.0	10.3	3.1	19855PG32T2

* This example study has two references.
** The first 5 digits refer to the file name, PG to pages, C to column, P to page, L to line, and T to Table.

from reference-based to study-based format.[37] We have found no other estimates of duration of the switch for other groups in health care. This may seem discouraging but the initial shift is a matter of days. Fully utilizing functionality does, however, take longer. Any initial investment is offset by the recognition of the daily waste occurring in the systematic reviewing process in discarding efforts of reviewers when supplying data to the next generation of researchers. The past and considerable efforts in concatenation into study and data extraction are routinely not re-issued. Repeating such efforts of merging (sometimes) hundreds of references into one study is not time-efficient. This merging and extraction of data should, of course, be entirely transparent to future users (study records can potentially hold such tracking information).

As already discussed, a relational database such as study-based register has the capacity to provide one-to-many relationships so that one study with many reports, but, in addition, conversely, one report referring to data from many studies could be linked and managed easily (many-to-one relationships).

1. Developing a register
The development process of specialized registers is described in Table 2. For study-based registers, guidelines should clarify what fields are required.[38,39] As the study

Box 1. Reasons for assessment and concatenation of all references of one study

1. Determining relevance
Screening either references or even an individual full-text report of the references may not provide enough information to convince the reviewer that the study meets the inclusion criteria.[30] Although references are assessed, it is the study that is the unit of analysis.

2. Minimizing the over-counting
Treating data from different references of the same study as separate studies will result in over-counting and spurious results.[29, 31, 32]

3. Minimizing the under-counting
Using data from only a selection of references of the study may fail to identify important outcomes in other relevant references of that same study. For example protocols of studies may report use of many more outcome measures than are finally reported in any one reference.[20]

record represents the total research activity in the project so data may have to be gleaned from different source documents. For example, one reference may record clinical outcomes of a study. A second reference may record economic data missing from the first reference – but all are brought together in greater or lesser detail in the 'parent' study record.[40]

Those undertaking systematic reviews already extract qualitative and quantitative information from all relevant references into a study record within their review. These data are often structured, use PICO headings and employ some sort of controlled vocabulary. These data can form the basis of a study record within a register for use by others interested in the area and maintenance continues as often study records have to be merged. Electronic study-register systems often automatically create a study record for every reference imported.[41] This one-to-one relationship is an understandable default but is inaccurate

Table 2. Levels of specialized registers

Beginners' level or reference-based register
a. Setting the scope of register
b. Developing, running, documenting, and saving the search strategies
c. De-duplicating and curettage the search results
d. Screening the search results
e. Developing minimum dataset for each reference record
f. Importing the reference records
g. Maintaining the reference records
h. Locating the full texts of all reference records
Intermediate level or study-based register
a. Developing coding scope and guideline for studies
b. Coding or extracting general data from each reference
c. Concatenation, merging, and cleaning the study records
d. Maintaining and updating the study records
Advanced level or automated study-based register
a. Classifying the studies under each review title ready to be done
b. Developing a machine-readable dataset for each study
c. Extracting all data from each study

and some merging of records will be necessary. This is not usually deleting one record in favor of another, but often the true merging of records to gain the most accurate description of the overall study. If, for example, that third reference of a recognizable study is the economics paper, and the package has erroneously considered that one reference to be a new unique study, it is important that (*i*) the economics outcomes are reported in the 'parent' study record; (*ii*) other outcomes are not deleted in favor of only economic outcomes; and that (*iii*) the unneeded study record is then deleted but the economics paper's reference record incorporated into the list of citations to that study.

Recognizing references to be of a single study is a skill. Usually, references with the same start time, locations, interventions and a number of participants are identified as references of the same study – although this is not always the case.[29] Recognition of single studies is assisted by use of study acronyms or trial registry number but these, unfortunately, remain the exception rather than the rule. Some software assists this process by pattern recognition within reference records for existing studies. For example, machines can recognize if, in a new reference, authors are identical to those in an existing study, publication dates are very close, interventions and numbers of participants are the same. An Information Specialist on top of her/his topic area quickly gains skills in study recognition. Furthermore, those using references and studies within reviews have to carefully scrutinize all references. Their view of what constitutes a study can be invaluable, save much time and be recycled into the Information Specialist's register.

If all steps are followed based on documented guidelines, it is easy to organize studies into review clusters. For instance, if there are 11 studies comparing 'Intervention A' versus 'Intervention B' for 'Condition C', as coded in fields of study record, all these studies could be listed under the potential review title 'Intervention A versus intervention B for condition C'. This review cluster is appended to with incoming relevant studies identified by the Information Specialist and, thereafter, end reviewers need to undertake little or no additional effort searching or screening.

If the full study data extraction is already undertaken, and this is stored within a higher level of study-based register, each study record will consist of a dataset in tabulated format (for readers) and XML-tagged machine-readable format (for software in which each piece of data/information has been linked to its original specific site within the source document). Since we could not identify any structure to cover a study record in the systematic review, Supplementary file 1 demonstrates a proposed XML structure to store extracted study data. A standard tagged and structured machine-readable framework can store data right down to the individual patient data level. Structuring for machine-readablity requires knowledge of the needs of both reviewers and computer scientists and a

malleable design to be open to future developments. Much has already been documented regarding such efforts for projects such as Distiller SR, Covidence, and Systematic Review Data Repository (SRDR).[5,42]

Challenges
1. The concept
The currency of the Information Specialist in health care has, for so long, been the individual reference (1 record for 1 reference) and making the jump to considering studies (1 record to potentially many references of 1 study) as the primary unit of information can be a conceptual challenge. After all, even cursory use of databases such as MEDLINE can cause nagging discomfort that information is being identified for reviewers that is less ordered and sorted than is ideal. For example, searching for CATIE, a large trial acronym, in the title field of MEDLINE will reveal multiple references (170 references at the time of writing). More latterly, solutions to this discomfort been conceived and so the issue is more acknowledged.[20,29-31] The process of systematic reviewing does help make this conceptual jump with the use of 'study tags' under which all relevant references are listed.

In a physical library with books, a librarian undertakes collection development, cataloging, classification and dissemination of information. A reference-based register with just collection development and no coding is the equivalent of maintaining a stack of books with no classification. Whilst 'save the time of the reader' is one of Five Laws of Library Science[43] 'save the time of the reviewer' is, we argue, best addressed in a study-based register. Also having a reference-based register 'might work' for some research groups, however, study-based register 'might work even better' saving the time and money for both the Information Specialists and systematic reviewers.

2. Responsibility
In reality, the process of identification of studies is a shared responsibility between Information Specialist and reviewer. Each has different skills to bring to the process. The former has competency in the identification of records, and knowledge of what each record should contain. For example, with reference records indexed with study identifiers (e.g., trial registry number), the Information Specialist is in a pivotal position to help the reviewer avoid needless effort linking references to a study. The reviewer, however, having inspected the detail of each reference, should be able to supply an authoritative study record back to the Information Specialist for their use or for the next reviewer.

Even now the responsibility for concatenation is shared. However, study-based registers allow this responsibility to be undertaken more easily by the person who maintains the register. The nature of systematic reviewing is that everything is double checked. By an iterative process, the study record evolves to be an increasingly accurate report of the primary investigation. To continually pass this responsibility down the line to reviewers creates an unnecessary waste and opportunity for inaccuracy.

3. Practicalities
For the gains which we outline above, development of such a study-based register involves an investment of effort – some pain - often from Information Specialists. Currently, Information Specialists working in systematic reviewing of health care interventions are busy and the thought of further work and/or responsibility may be unwelcome. However, much current work is inefficient and we argue the work creating a study-based register is an investment. The 'pain' involved makes the role of Information Specialist much more sophisticated and prepares that person to manage a register suitable for the data needs of the 21st century.

Coding of study happens by a shared iterative process as outlined above. Many reference records already contain study codes which can be imported into the study record. For example, the PT field of MEDLINE may contain "Randomized Controlled Trial" – a methodological term relating to the study, or the SI field contains the International Standard Randomized Controlled Trial Number (ISRCTN) – again data that relates to the study rather than the reference. These data can be imported automatically into the study record at no cost of time. As reviewers 'use' the study, adding to the complexity but also the utility of the record, more sophisticated data from this investment of effort can be curated and stored ready for the next reviewer.

4. Software
Although there are online clinical trials register such as ClinicalTrials.Gov, ISRCTN, and WHO International Clinical Trials Registry Platform (including 16 trial registers), however, none of these are study-based registers and none are aimed to support systematic reviews and provide a very limited number of fields for each study record. Even before the initiation of these sources, in 1995, Cochrane Stroke Group started using study-based register and mental health followed.[44-49] A year after Stroke Group, UK Cochrane Centre supported MeerKat Working Group to develop a study-based register system.[23,50] In 2003, there were 10 Cochrane groups using MeerKat and 5 other groups considering its use.[23] By 2005, at least 12 Cochrane groups (out of 37 responses) were maintaining a study-based register on MeerKat, ProCite, Reference Manager or RefTrak.[25,51-53] In 2008, Cochrane started developing a new program, Cochrane Register of Studies (CRS), to be used by all groups.[54-59] A survey in 2014 showed that 8 out of 29 respondent groups are using a study-based register to some extent.[12]

There are several reference management programs and some study-based register programs. Despite the

decision of Cochrane to move to CRS as the only program for managing information on both references and studies, groups do tend to use other programs as well as CRS.[8,9,14,44,60-75] Although others have pioneered software for study-based registers[44,45] there are few packages that are generic, accessible, or customizable. EndNote, ProCite, and Reference Manager are popular reference management programs that can be modified to include some features of study-based registers. These bibliographic packages do not lend themselves easily to this adaption and relational databases have considerable advantages. Since the purpose of programs such as EndNote or Reference Manager is not primarily managing studies for systematic reviews, using them in a study-based fashion requires time and effort – and there is an element of fitting a 'square-peg-in-the-round-hole'. Tailor-made relational programs such as MeerKat or RefTrak are much better and potentially more malleable. CRS continues to evolve.

No matter what software or application is used, the transition to a study-based register necessitates the acquisition of a skill-set to use the package to full capacity from the study-perspective. This transition is helped by practice, training and mentoring.

5. Data ownership
As mentioned before, there are other resources, such as online registries of trials, that do provide some useful data for studies; however, there are certain barriers in terms of copyright and legal issues such as 'who owns the data'.[76] Such limitations make it hard to use and share the data openly and import or share them from other resources in study-based registers.

6. Study designs
Since the current development of study-based registers in Cochrane, is mainly focused on RCTs/CCTs, and because we work within that organization, we have not discussed involving other study designs. There is, however, a possibility to include all empirical study designs in study-based registers of the future. This may cause new challenges for the development of the registers involving different sets of meta-data. Some Cochrane groups are already considering involving more diverse study designs in their reviews. However, these other study designs are not a priority for consideration in the technological development of specialized registers. The development of such registers has made it possible to link the studies from pre-clinical sciences to clinical sciences. Such link could reduce the waste by avoiding a clinical trial research where a systematic review of pre-clinical studies about ineffectiveness of an intervention exists.[77]

Gains
Having a study-based register gives review groups the advantage of knowing with some accuracy how many systematic reviews (or at least comparisons) there are to

---

Methodology Highlights

What is current knowledge?

√ A systematic review is required before starting each biomedical research.

√ Doing a systematic review is time-consuming, costly and requires training.

√ To do a systematic review, the review team should search relevant databases with suitable search strategy, de-duplicate the results, screen the results, obtain the full texts, check the papers against the eligibility criteria of the review, collect the papers relevant to one study under one study name (studifying) and then extract and analyze the data.

What is new here?

√ Study-based registers could link the clinical and pre-clinical studies related to the same research question to support the translational research.

√ Study-based registers are saving the time and cost of systematic review for the team skipping the searching, de-duplicating, screening, finding the full texts, criteria checking and studifying steps. The systematic review could start with data extraction or meta-analysis.

√ There are three different levels of specialized registers which are now the milestone of automation of systematic reviews.

---

cover in a topic area – and to more accurately estimate the future work-load.[78] Such knowledge allows accurate and efficient prioritization of effort.

Current resources, if used effectively, are sufficient to increase productivity. Using the study as the currency of the Information Specialist, modifying that role within a review group to create and maintain the study-based register, and finally using such a register will greatly increase the pace and efficiently of information exchange.

Information Specialists could then undertake regular – perhaps semi-automated – searches, screen, and code and include relevant records within the study-based register. New references of existing studies would be added to the existing record and, if this study is used within an existing review, the reviewers would be notified. New studies relevant to existing review topics would also cause reviewers to be alerted. Rather than passively waiting for reviewers to request update searches Information Specialists would be proactive in helping the update. From the ordered study-based register studies can be linked to topics and be instantly ready for new reviewers. Should data from a study have been extracted for use in an existing review, these detailed data could be supplied from a sophisticated study-based register in an appropriate format to anyone undertaking a new review necessitating the use of the same study.

Conclusions
The ultimate goal of having a study-based register is to facilitate efficient production of systematic reviews

providing rapid but accurate evidence for decision-makers. The future will involve increasing automation. Document optimization now allows much more reliable auto-data extraction.[79] Programs already exist for semi-automated data extraction from randomized trials[5] and text mining techniques are increasingly sophisticated. The automatic synthesis of data is not far away, perhaps driven by users' needs rather than those of policymakers. Limited automatic write-up of synthesized evidence already exists.[79-81] These next years will see a swift synthesis of best and personalized evidence of the effects of health care in the hands of anyone. At the heart of this exciting prospect should be the role of the Information Specialist – but a role fit for the 21st century and not one that is dated and wasteful. Moving from reference-based register toward study-based register is, we think, inevitable. The infrastructures are ready for such movement.

Ethical issues
There is none to be declared.

Competing interests
No competing interests to be disclosed.

Acknowledgements

Supplementary Materials
Supplementary file 1 contains a Randomized Controlled Trial Study in XML Format.

References
1. O'Reilly CA. The use of information in organizational decision making: A model and some propositions. *Res Organ Behav* **1983**;5:103-39.
2. Gonzalez A. PNA/08-02: "Development and implementation of a methodology to discover knowledge in corporative database as support to the organizational making decision process". CUBAENERGIA; 2002.
3. Ingwersen P, Järvelin K. Evaluation Measures: Generalized Recall and Precision. In: *The Turn Integration of Information Seeking and Retrieval in Context*. Dordrecht, The Netherlands: Springer; **2005**. p. 180-1.
4. The Cochrane Collaboration. The Cochrane Collaboration. 2015. http://www.cochrane.org/.
5. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. *Syst Rev* **2014**;3:74. doi: 10.1186/2046-4053-3-74.
6. Tsafnat G, Dunn A, Glasziou P, Coiera E. The automation of systematic reviews. *BMJ* **2013**; 346: f139. doi: 10.1136/bmj.f139.
7. Higgins G, ed. *Register & handsearch submissions for CENTRAL: all you need is trials*! Freiburg, Germany: Cochrane Colloquium; **2008**.
8. Manheimer E, Higgins G, Wieland S, eds. *Using ProCite for creating and submitting trials databases to CENTRAL*. Melbourne, Australia: Cochrane Colloquium; **2005**.
9. Manheimer E, Lefebvre C, Timimi H, Rutks I, Ghersi D, edis. Lessons Learned and Ongoing Challenges in Developing an International, Collaborative Register of Controlled Trials with Many Contributors. Stavanger, Norway: Cochrane Qolloquium; **2013**.
10. The Cochrane Collaboration. About The Cochrane Library. http://wwwthecochranelibrarycom/view/0/AboutTheCochraneLibraryhtml. 2015.
11. Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* **2002**;25:38-64. doi: 10.1177/016327870202500104.
12. Noel-Storr A. Specialised registers: a snapshot of activity. 2014.
13. Thompson E, Macbeth F. New Directions in the Evaluation and Presentation of Clinical Research in Lung Cancer. In: Jeremic B, ed. Advances in Radiation Oncology in Lung Cancer. Springer: Berlin; **2011**. p. 505-12. doi: 10.1007/978-3-642-19925-7.
14. Busgeeth K, Siegfried N, Fenton M, Adams C, Clarke M, Eisinga A, et al, editors. Improving our quality by developing the HIV/AIDS Specialized Register: experiences and lessons from the inaugural Cochrane Collaboration Visiting Fellowship. Melbourne, Australia: Cochrane Colloquium; **2005**.
15. Schneck C, Roberts S, Adams CE. Ferret reborn: library holdings linkage to CRS. Cochrane Schizophrenia Group; 2013.
16. Bashir R, Bourgeois FT, Dunn AG. A systematic review of the processes used to link clinical trial registrations to their published results. *Syst Rev* **2017**; 6: 123. doi: 10.1186/s13643-017-0518-3.
17. CRS Team. Studification: maintaining your study based register. CRS Webinar; **2013**.
18. Ebrahim S, Montoya L, Kamal El Din M, Sohani ZN, Agarwal A, Bance S, et al. Randomized trials are frequently fragmented in multiple secondary publications. *J Clin Epidemiol* **2016**;79:130-9. doi: 10.1016/j.jclinepi.2016.05.016.
19. Fattahi R. Super records: An approach towards the description of works appearing in various manifestations. *Libr Rev* **1996**;45:19-29. doi: 10.1108/eum0000000004129.
20. Altman D, Furberg C, Grimshaw J, Shanahan D. Linked publications from a single trial: a thread of evidence. *Trials* **2014**; 15:369.
21. Durao S, Volmink J, van Binsbergen J, editors. Proposal to establish a Cochrane Diet and Nutrition Trials Register. Québec, Canada: Cochrane Colloquium; **2013**.
22. Shokraneh F. PICO Framework: Two Decades of Variation and Application. Evidence Live; 22-24 June 2016; University of Oxford, Oxford, UK.
23. Kaur S, de Souza M, Fenton M, Adams CE. MeerKat 1.4 – and beyond. Cochrane Schizophrenia Group; 2009.
24. MeerKat Working G, editor. MeerKat: getting the most out of your study-based Specialized Register. Melbourne, Australia: Cochrane Colloquium; **2005**.
25. Wright J. MeerKat User Guide For MeerKat Version 1.4. Grant T, editor. Leeds: Cochrane Schizophrenia Group; **2006**.
26. Stoelwinder J, Hill S, Oliver S, Broclain D, Wensing M, Brewer L, editors. Coding Scheme for the Cochrane Consumers and Communication Review Group (CCCRG) Specialised Register. Cochrane Colloquium; **2003**; Stavanger, Norway.
27. McDonald S, Thomas J, Wallace S, Elliott J, editors. Shortening the pipeline: the use of data mining to link new trials to Cochrane Reviews. Québec, Canada: Cochrane Qolloquium; **2013**.
28. Shokraneh F, Adams CE. Increasing value and reducing waste in data extraction for systematic reviews: tracking data in data extraction forms. *Syst Rev* **2017**;6:153. doi: 10.1186/s13643-017-0546-z.
29. Huston P, Moher D. Redundancy, disaggregation, and the integrity of medical research. *Lancet* **1996**;347:1024-6.
30. Leizorovicz A, Haugh M, Boissel JP. Meta-analysis and multiple publication of clinical trial reports. *Lancet* **1992**;340:1102-3.
31. Tramer MR, Reynolds DJM, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. *BMJ* **1997**;315:635-40. doi: 10.1136/bmj.315.7109.635.
32. Senn SJ. Overstating the evidence:double counting in meta-analysis and related problems. *BMC Med Res Methodol* **2009**;9:10. doi: 10.1186/1471-2288-9-10
33. Foxlee R, Noel-Storr A, Dooley G, Littlewood A, Salzwedel D, editors. Data cleaning in the Cochrane Register of Studies. Québec, Canada: Cochrane Qolloquium; **2013**.
34. Dooley G, Anstee D, Foxlee R, editors. Link checking and cleaning:preparing records for Linked Data. Cochrane Hyderabad, India: Colloquium; **2014**.
35. Noel-Storr A, Foxlee R, Dooley G, Littlewood A, Salzwedel D, editors. Harmonising Cochrane content in the CRS:moving towards a nimble global dataset. Cochrane Hyderabad, India: Colloquium;

2014.

36. Ingwersen P, Järvelin K. Research Methods. In: The Turn Integration of Information Seeking and Retrieval in Context. Dordrecht, The Netherlands: Springer; **2005**. p. 9.

37. Hermans D, McShane R, Schneider LS. The Cochrane Dementia Group's study-based register of prevention trials. AAlzheimers Dement (N Y) **2007**;3:S125.

38. Noel-Storr A. Standardisation: study-based meta data. CRS Webinar; **2013**.

39. Noel-Storr A. Standardisation: bibliographic data in the CRS. CRS Webinar; **2013**.

40. Noel-Storr A, Dooley G, editors. Studification in the Cochrane Register of Studies. Hyderabad, India: Cochrane Qolloquium; **2014**.

41. Shokraneh F, Adams CE. Potentials of Registers of Randomized Controlled Trials in Automating the Systematic Reviews. Edinburgh, Scotland, UK: REWARD EQUATOR; **2015**.

42. Li T, Vedula SS, Hadar N, Parkin C, Lau J, Dickersin K. Innovations in data collection, management, and archiving for systematic reviews. *Ann Intern Med* **2015**;162:287-94. doi: 10.7326/m14-1603.

43. Ranganathan SR. The Five Laws of Library Science:Madras. London, Edward Goldston, Ltd: The Madras Library Association; **1931**.

44. Cochrane Stroke Group. DORIS. Cochrane Stroke Group; 2015. p. http://stroke.cochrane.org/doris.

45. Fraser H, Mclnnes A, Carratt G, Counsell C, editors. RefTraK - The Cochrane Stroke Group's Specialised Register Management System:A Powerful Tool for Supporting Reviewers. Rome, Italy: Cohrane Colloquium; **1999**.

46. Portfors P, Wahlbeck K, Tuunainen A, Ekqvist M, Heino T, Englund J, et al., editors. Dissemination of Cochrane Mental Health Groups' specialised registers on the web - the EU-PSI project and the PsiTri register. Cologne, Germany: European Association for Health Information and Libraries Conference; **2002**.

47. Adams C, Stancliffe R, editors. The EU Mental Health Library. The Netherlands, Amsterdam: Elsevier Science; **2002**.

48. Wahlbeck K, Adams C, Birks J, Churchill R, Ferri M, Gilbody S, et al, editors. The EU-(Project:Dissemination of mental health trial evidence. Cape Town, South Africa: Cochrane Colloquium; **2000**.

49. Thomas B, Mclnnes A, editors. The Cochrane Stroke Group's Specialised Register of Trials: A Comprehensive and Unique Resource for Reviewers. Rome, Italy: Cochrane Colloquium; **1999**.

50. MeerKat Working Group, editor. MeerKat:getting the most out of your study-based Specialized Register. Melbourne, Australia: Cochrane Colloquium; **2005**.

51. Wright J, editor. Study-based registers. Dublin, Ireland: Cochrane Colloquium; **2006**.

52. Wright J. Study-based registers – a brief introduction. Cochrane Schizophrenia Group; 2006.

53. Wright J. Setting up and maintaining a study-based register. Cochrane Schizophrenia Group; 2006.

54. The Cochrane Collaboration. Cochrane Register of Studies (CRS). The Cochrane Collaboration; 2015. http://www.metaxis.com/CRSSoftwarePortal/Index.asp.

55. Foxlee R, editor. The development of the new Cochrane Register of Controlled Studies. Freiburg, Germany: Cochrane Qolloquium; **2008**.

56. Foxlee R, Dooley G, Jones L. Development of the Cochrane Register of Studies. *Cochrane Database Syst Rev* **2010**;156.

57. Foxlee R, Dooley G, Jones L. User testing of the Cochrane Register of Studies. *Cochrane Database Syst Rev* **2010**;162-3.

58. Foxlee R, Fiander M, Dooley G, Salzwedel D, editors. Using the Cochrane Register of Studies to studify your Specialised Register. Auckland, New Zealand: Cochrane Qolloquium; **2012**.

59. Malouf R, Noel-Storr A, Collins H, McShane R, Schneider LS. ALOIS: Alzheimer's and cognitive improvement studies register, a free, on-line register of dementia and cognitive enhancement trials (http://www.medicine.ox.ac.uk/alois/). *Alzheimers Dement (N Y)* **2009**;5:P251-2. doi: 10.1016/j.jalz.2009.04.264.

60. Cochrane Renal Group. Locating studies for your systematic review. Cochrane Renal Group; 2005.

61. Higgins G. The Cochrane Haematological Malignancies Group Specialised Register. RundBrief **2003**;7-8.

62. Cochrane Schizophrenia Group. CSzG Specialised Register. Cochrane Schizophrenia Group; 2014.

63. Adams CE, Coutinho ES, Davis J, Duggan L, Leucht S, Li C, et al. Cochrane Schizophrenia Group. *Schizophr Bull* **2008**;34:259-65. doi: 10.1093/schbul/sbm167.

64. Hovhannisyan K, Higgins G, Hampson L, Wallace S. Trials Search Co-ordinators, Archie & RevMan 5. T*Cochrane Database Syst Rev* **2010**;145.

65. Zani B, Siegfried N, Pienaar E, editors. Randomized controlled trials of HIV/AIDS prevention and treatment in Africa:results from the Cochrane HIV/AIDS Specialized Register. Singapore: Cochrane Colloquium; **2009**.

66. Zani B, Oliver J, Siegfried N, editors. Sorting the wheat from the chaff:How does the Cochrane HIV/AIDS Specialized Trials Register compare to searching standard electronic databases? Madrid, Spain: Cochrane Qolloquium; **2011**.

67. Monalisa J, Roberts S, Adams CE, editors. The Cochrane Schizophrenia Group (CSzG) Specialised Register. 3rd South Asian Regional Symposium on Evidence Informed Healthcare; **2010**; Vellore, India.

68. Pienaar E, Siegfried N, Oliver J, Zani B, editors. The Cochrane HIV/AIDS Trials Register:five years of achievements, progress and challenges. Singapore: Cochrane Colloquium; **2009**.

69. Henderson S, Hampson L, Neilson J. Cochrane Pregnancy and Childbirth Group: process and achievements. German Journal for Evidence and Quality in Health Care **2002**;96:687-8.

70. Cochrane Pregnancy and Childbirth Group. Pregnancy and Childbirth Group's Trials Register. "Cochrane Pregnancy and Childbirth Group"; 2015. http://pregnancy.cochrane.org/pregnancy-and-childbirth-groups-trials-register.

71. Cochrane Dementia and Cognitive Improvement Group. Resources for review authors. Cochrane Dementia and Cognitive Improvement Group; 2014.

72. Noel-Storr A. ALOIS: a comprehensive register of demential studies. 2011. http://www.medicine.ox.ac.uk/ALOIS/.

73. Noel-Storr A, Malouf R, editors. Launch of ALOIS - an online study-based register of research from the Cochrane group. 4th UK Dementia Congress; **2009**; Harrogate, UK.

74. Noel-Storr A, McShane R. ALOIS-the next phase:a comprehensive study-based register of diagnostic studies (www. medicine. ox. ac. uk/alois). *Alzheimers Dement (N Y)* **2011**;7:S324-S5. doi: 10.1016/j.jalz.2011.05.943.

75. Ssemanda E, Dickersin K, Li T, Scherer R, Ervin A, Hawkins B, editors. The E-trials project:first steps in the development of a study-based eyes and vision trials database. Freiburg, Germany Cochrane Colloquium; **2008**.

76. Bierer BE, Crosas M, Pierce HH. Data Authorship as an Incentive to Data Sharing. *N Engl J Med* **2017**;376:1684-7. doi: 10.1056/NEJMsb1616595.

77. Glasziou P, Chalmers I. How systematic reviews can reduce waste in research. **2015**. http://blogs.bmj.com/bmj/2015/10/29/how-systematic-reviews-can-reduce-waste-in-research/. Accessed September 11, 2017

78. McClenaghan E, Roberts S, Adams CE. So – how many reviews do you have to do? Cochrane Schizophrenia Group; 2012.

79. Torres MM, Rana A, Stark B, Hagmann S, Knahl C, Polzmacher S, et al. Automatize or perish: Towards user-led, instant and (at least) semi-automated systematic reviews. Evidence Live; 13-14 April 2015; University of Oxford, Oxford, UK.

80. Adams CE, Polzmacher S, Wolff A. Systematic reviews: work that needs to be done and not to be done. *J Evid Based Med* **2013**;6:232-5. doi: 10.1111/jebm.12072.

81. Torres MM, Rana A, Stark B, Hagmann S, Knahl C, Polzmacher S, et al. The (pregnant) elephant in the room:a tool to speed up delivery of Cochrane Reviews. Evidence Live; 13-14 April 2015; University of Oxford, Oxford, UK.