What's new in Version 5.0.2 (September 2009)

Main reason for update: to replicate the revised reprint of the book (John Wiley & Sons, 2009).

Chapter 6

- Boxes 6.4.a, 6.4.b, 6.4.c, 6.4.d : last lines of each strategy corrected.

Chapter 13

- Table 13.2.b: entries for ITS and ChBA have been corrected for the first two rows.

Chapter 6: Searching for studies

Authors: Carol Lefebvre, Eric Manheimer and Julie Glanville on behalf of the Cochrane Information Retrieval Methods Group.

This extract is from *Handbook* version 5.0.2. For guidance on how to cite it, see Section 6.7. The material is also published in Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions* (ISBN 978-0470057964) by John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, Telephone (+44) 1243 779777; Email (for orders and customer service enquiries): cs-books@wiley.co.uk. Visit their Home Page on www.wiley.com.

Key points

- Review authors should work closely from the start with the Trials Search Co-ordinator (TSC) of their Cochrane Review Group (CRG).
- Studies (not reports of studies) are included in Cochrane reviews but identifying reports of studies is currently the most convenient approach to identifying the majority of studies and obtaining information about them and their results.
- Trials registers and trials results registers are an increasingly important source of information.
- The Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and EMBASE (if access is available to either the review author or TSC) should be searched for all Cochrane reviews, either directly or via the CRG's Specialized Register.
- Searches should seek high sensitivity, which may result in relatively low precision.
- Too many *different* search concepts should be avoided, but a wide variety of search terms should be combined with OR within *each* concept.
- Both free-text and subject headings should be used (for example Medical Subject Headings (MeSH) and EMTREE).
- Existing highly sensitive search strategies (filters) to identify randomized trials should be used, such as the newly revised Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE (but do not apply these filters in CENTRAL).

6.1 Introduction

Cochrane Review Groups (CRGs) are responsible for providing review authors with references to studies that are possibly relevant to their review. The majority of CRGs employ a dedicated Trials Search Co-ordinator to provide this service (see Section 6.1.1.1). The information in this chapter is designed to assist authors wishing to undertake supplementary searches for studies and to provide background information so that they can better understand the search process. In all cases review authors should contact the Trials Search Co-ordinator of their CRG before starting to search, in order to find out the level of support they provide.

This chapter will also be useful to Trials Search Co-ordinators who are new to their post, as well those who are more experienced, who may wish to consult this chapter as a reference source.

This chapter outlines some general issues in searching for studies; describes the main sources of potential studies; and discusses how to plan the search process, design and carry out search strategies, manage references found during the search process and correctly document and report the search process.

This chapter concentrates on searching for randomized trials. Many of the search principles discussed, however, will also apply to other study designs as discussed elsewhere. For some review topics, for example complex interventions, it may be necessary to adopt other approaches and to include studies other than randomized trials. Review authors are recommended to seek specific guidance from their CRG and refer also to the relevant chapters of this *Handbook*, such as Chapter 13 for non-randomized studies, Chapter 14 for adverse effects, Chapter 15 for economics data, Chapter 17 for patient-reported outcomes, Chapter 20 for qualitative research and Chapter 21 for reviews in health promotion and public health. Review authors searching for studies for inclusion in Cochrane reviews of diagnostic test accuracy should refer to the *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy*.

The numerous web sites listed in this chapter were checked in June 2008.

6.1.1 General issues

6.1.1.1 Role of the Trials Search Co-ordinator

The Trials Search Co-ordinator for each CRG is responsible for providing assistance to authors with searching for studies for inclusion in their reviews. The range of assistance varies according to the resources available to individual CRGs but may include some or all of the following: providing relevant studies from the CRG's Specialized Register (see Section 6.3.2.4 for more detail), designing search strategies for the main bibliographic databases, running these searches in databases available to the CRG, saving search results and sending them to authors, advising authors on how to run searches in other databases and how to download results into their reference management software (see Section 6.5). Contact your Trials Search Co-ordinator before you start searching to find out the level of assistance offered.

If a CRG is currently without a Trials Search Co-ordinator authors should seek the guidance of a local healthcare librarian or information specialist, where possible one with experience of conducting searches for systematic reviews.

6.1.1.2 Minimizing bias

Systematic reviews of interventions require a thorough, objective and reproducible search of a range of sources to identify as many relevant studies as possible (within resource limits). This is a major factor in distinguishing systematic reviews from traditional narrative reviews and helps to minimize bias and therefore assist in achieving reliable estimates of effects.

A search of MEDLINE alone is not considered adequate. A systematic review showed that only 30% - 80% of all known published randomized trials were identifiable using MEDLINE (depending on the area or specific question) (Dickersin 1994). Even if relevant records are in MEDLINE, it can be difficult to retrieve them (Golder 2006, Whiting 2008). Going beyond MEDLINE is important not only for ensuring that as many relevant studies as possible are identified but also to minimize selection bias for those that are found. Relying exclusively on a MEDLINE search may retrieve a set of reports unrepresentative of all reports that would have been identified through a comprehensive search of several sources.

Time and budget restraints require the review author to balance the thoroughness of the search with efficiency in use of time and funds and the best way of achieving this balance is to be aware of, and try to minimize, the biases such as publication bias and language bias that can result from restricting searches in different ways (see Chapter 10, Section 10.2).

6.1.1.3 Studies versus reports of studies

Systematic reviews have studies as the primary units of interest and analysis. However, a single study may have more than one report about it and each of these reports may contribute useful information for the review (see Chapter 7, Section 7.2). For most of the sources listed in Section 6.2, the search process will retrieve individual reports of studies, however there are some study-based resources, such as trials registers and trials results databases (see Sections 6.2.3.1 to 6.2.3.4).

6.1.1.4 Copyright and licensing

It is Cochrane Collaboration policy that all review authors and others involved in the Collaboration should adhere to copyright legislation and the terms of database licensing agreements. With respect to searching for studies, this refers in particular to adhering to the terms and conditions of use when searching databases and downloading records and adhering to copyright legislation when obtaining copies of articles. Review authors should seek guidance on this from their Trials Search Co-ordinator or local healthcare librarian, as copyright legislation varies across jurisdictions and licensing agreements across institutions.

6.1.2 Summary points

- Cochrane review authors should seek advice from the Trials Search Co-ordinator of their Cochrane Review Group (CRG) *before* starting a search.
- If the CRG is currently without a Trials Search Co-ordinator, seek the guidance of a local healthcare librarian or information specialist, where possible one with experience of searching for systematic reviews.
- Use the Table of Contents to navigate to specific sections of this chapter.
- A search of MEDLINE alone is not considered adequate.
- It is Cochrane Collaboration policy that all review authors and others involved in the Collaboration should adhere to database licensing terms and conditions of use and copyright legislation.

6.2 Sources to search

6.2.1 Bibliographic databases

6.2.1.1 Bibliographic databases – general introduction

Searches of health-related bibliographic databases are generally the easiest and least time-consuming way to identify an initial set of relevant reports of studies. Some bibliographic databases, such as MEDLINE and EMBASE, include abstracts for the majority of recent records. A key advantage of these databases is that they can be searched electronically both for words in the title or abstract and by using the standardized indexing terms, or controlled vocabulary, assigned to each record (see Section 6.4.5).

The Cochrane Collaboration has been developing a database or register of reports of controlled trials called The Cochrane Central Register of Controlled Trials (CENTRAL). This is considered to be the

best single source of reports of trials that might be eligible for inclusion in Cochrane reviews. The three bibliographic databases generally considered to be the most important sources to search for reports of trials – CENTRAL, MEDLINE and EMBASE – are described in more detail in subsequent sections.

Databases are available to individuals for a fee, on a subscription or on a 'pay-as-you-go' basis. They can also be available free at the point of use through national provisions, site-wide licences at institutions such as universities or hospitals, through professional organizations as part of their membership packages or free of charge on the internet.

There are also a number of international initiatives to provide free or low-cost online access to databases (and full-text journals) over the internet. The Health InterNetwork Access to Research Initiative (HINARI) provides access to a wide range of databases including *The Cochrane Library* and nearly 4000 major journals from a wide range of publishers in biomedical and related social sciences, for healthcare professionals in local, not-for-profit institutions in over 100 low-income countries.

- www.who.int/hinari/en/

The International Network for the Availability of Scientific Publications (INASP) also provides access to a wide range of databases including *The Cochrane Library* and journals. Journal titles available vary by country. For further details see:

- www.inasp.info/file/68/about-inasp.html

Electronic Information for Libraries (eIFL) is a similar initiative based on library consortia to support affordable licensing of journals in 50 low-income and transition countries in central, eastern and south-east Europe, the former Soviet Union, Africa, the Middle-East and south-east Asia.

- www.eifl.net/cps/sections/about

For more detailed information about how to search these and other databases refer to Sections 6.3.3 and 6.4.

6.2.1.2 The Cochrane Central Register of Controlled Trials (CENTRAL)

The Cochrane Central Register of Controlled Trials (CENTRAL) serves as the most comprehensive source of reports of controlled trials. CENTRAL is published as part of *The Cochrane Library* and is updated quarterly. As of January 2008 (Issue 1, 2008), CENTRAL contains nearly 530,000 citations to reports of trials and other studies potentially eligible for inclusion in Cochrane reviews, of which 310,000 trial reports are from MEDLINE, 50,000 additional trial reports are from EMBASE and the remaining 170,000 are from other sources such as other databases and handsearching.

Many of the records in CENTRAL have been identified through systematic searches of MEDLINE and EMBASE, as described in Sections 6.3.2.1 and 6.3.2.2. CENTRAL, however, includes citations to reports of controlled trials that are not indexed in MEDLINE, EMBASE or other bibliographic databases; citations published in many languages; and citations that are available only in conference proceedings or other sources that are difficult to access (Dickersin 2002). It also includes records from trials registers and trials results registers (see Section 6.2.3).

CENTRAL is available free of charge to all CRGs through access to *The Cochrane Library*. The web address for *The Cochrane Library* is: http://www.thecochranelibrary.com. Many health and academic institutions and organizations provide access to their members, and in many countries there is free

access for the whole population (for example through funded national licences or arrangements for low-income countries). Information about access to *The Cochrane Library* for specific countries can be found under 'Access to Cochrane' at the top of *The Cochrane Library* home page.

6.2.1.3 MEDLINE and EMBASE

MEDLINE currently contains over 16 million references to journal articles from the 1950s onwards. Currently 5,200 journals in 37 languages are indexed for MEDLINE:

- www.nlm.nih.gov/pubs/factsheets/medline.html

PubMed provides access to a free version of MEDLINE that also includes up-to-date citations not yet indexed for MEDLINE:

- www.nlm.nih.gov/pubs/factsheets/pubmed.html

Additionally, PubMed includes records from journals that are not indexed for MEDLINE and records considered 'out-of-scope' from journals that are partially indexed for MEDLINE. For further information about the differences between MEDLINE and PubMed see:

- www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html

MEDLINE is also available on subscription from a number of online database vendors, such as Ovid. Access is usually free to members of the institutions paying the subscriptions (e.g. hospitals and universities).

The US National Library of Medicine (NLM) has developed the NLM Gateway, which allows users to search MEDLINE or PubMed together with other NLM resources simultaneously such as the Health Services Research Projects database (HSRProj), Meeting Abstracts and the TOXLINE Subset for toxicology citations.

- gateway.nlm.nih.gov/gw/Cmd

EMBASE currently contains over 12 million records from 1974 onwards. Currently

4,800 journals are indexed for EMBASE in 30 languages.

- www.info.embase.com/embase_suite/about/brochures/embase_fs.pdf

EMBASE.com is Elsevier's own version of EMBASE that, in addition to the 12 million EMBASE records from 1974 onwards, also includes over 7 million unique records from MEDLINE from 1966 to date, thus allowing both databases to be searched simultaneously.

- www.info.embase.com/embase_com/about/index.shtml

In 2007, Elsevier launched EMBASE Classic which now provides access to records digitized from the Excerpta Medica print journals (the original print indexes from which EMBASE was created) from 1947 to 1973.

- www.info.embaseclassic.com/pdfs/factsheet.pdf

EMBASE is only available by subscription. Authors should check if their CRG has access and, if not, whether it is available through their local institution's library.

For guidance on how to search MEDLINE and EMBASE for reports of trials, see Sections 6.3.3.2, 6.4.11.1 and 6.4.11.2 respectively.

Database overlap

Of the 4,800 journals indexed in EMBASE, 1,800 are not indexed in MEDLINE. Similarly, of the 5,200 journals indexed in MEDLINE, 1,800 are not indexed in EMBASE.

- www.info.embase.com/embase_suite/about/brochures/embase_fs.pdf

The actual degree of reference overlap varies widely according to the topic but studies comparing searches of the two databases have generally concluded that a comprehensive search requires that both databases be searched (Suarez-Almazor 2000). Although MEDLINE and EMBASE searches tend not to identify the same sets of references, they have been found to return similar numbers of relevant references.

6.2.1.4 National and regional databases

In addition to MEDLINE and EMBASE, which are generally considered to be the key international general healthcare databases, many countries and regions produce electronic bibliographic databases that concentrate on the literature produced in those regions, and which often include journals and other literature not indexed elsewhere. Access to many of these databases is available free of charge on the internet. Others are only available by subscription or on a 'pay-as-you-go' basis. Indexing complexity and consistency varies, as does the sophistication of the search interface, but they can be an important source of additional studies from journals not indexed in other international databases such as MEDLINE or EMBASE. Some examples are included in Box 6.2.a.

Box 6.2.a: Examples of regional electronic bibliographic databases

Africa: African Index Medicus

- indexmedicus.afro.who.int/

Australia: Australasian Medical Index (fee-based)

- www.nla.gov.au/ami/

China: Chinese Biomedical Literature Database (CBM) (in Chinese)

- www.imicams.ac.cn/cbm/index.asp

Eastern Mediterranean: Index Medicus for the Eastern Mediterranean Region

- www.emro.who.int/his/vhsl/

Europe: PASCAL (fee-based)

- international.inist.fr/article21.html

India: IndMED

- indmed.nic.in/

Korea: KoreaMed

- www.koreamed.org/SearchBasic.php

Latin America and the Caribbean: LILACS

- bases.bireme.br/cgi-bin/wxislind.exe/iah/online/?IsisScript=iah/iah.xis&base=LILACS&lang=i&Form=F

South-East Asia: Index Medicus for the South-East Asia Region (IMSEAR)

- library.searo.who.int/modules.php?op=modload&name=websis&file=imsear

Ukraine and the Russian Federation: Panteleimon

- www.panteleimon.org/maine.php3

Western Pacific: Western Pacific Region Index Medicus (WPRIM)

- wprim.wpro.who.int/SearchBasic.php

6.2.1.5 Subject-specific databases

Which subject-specific databases to search in addition to CENTRAL, MEDLINE and EMBASE will be influenced by the topic of the review, access to specific databases and budget considerations. Most of the main subject-specific databases are available only on a subscription or 'pay-as-you-go' basis. Access to databases is therefore likely to be limited to those databases that are available to the Trials Search Co-ordinator at the CRG editorial base and those that are available at the institutions of the review authors. A selection of the main subject-specific databases that are more likely to be available through institutional subscriptions (and therefore 'free at the point of use') or are available free of charge on the internet are listed in Box 6.2.b, together with web addresses for further information. Access details vary according to institution. Review authors should seek advice from their local healthcare librarian for access at their institution.

In addition to subject-specific databases, general search engines include:

- Google Scholar (free on the internet):
  - scholar.google.com/advanced_scholar_search?hl=en&lr=
- Intute (free on the internet):
  - www.intute.ac.uk/
- Turning Research into Practice (TRIP) database (evidence-based healthcare resource) (free on the internet):
  - www.tripdatabase.com/

Box 6.2.b: Examples of subject-specific electronic bibliographic databases

Biology and pharmacology

- Biological Abstracts / BIOSIS Previews:
  - biosis.org/
- Derwent Drug File:
  - scientific.thomson.com/support/products/drugfile/
- International Pharmaceutical Abstracts:
  - scientific.thomson.com/products/ipa/

Health promotion

- BiblioMap – EPPI-Centre database of health promotion research (free on the internet):
  - eppi.ioe.ac.uk/webdatabases/Intro.aspx?ID=7
- Database of Promoting Health Effectiveness Reviews (DoPHER) (free on the internet):
  - eppi.ioe.ac.uk/webdatabases/Intro.aspx?ID=2

International health

- Global Health:
  - www.cabi.org/datapage.asp?iDocID=169

- POPLINE (reproductive health) (free on the internet):
    - db.jhuccp.org/ics-wpd/popweb/

Nursing and allied health

- Allied and Complementary Medicine (AMED):
    - www.bl.uk/collections/health/amed.html
- British Nursing Index (BNI):
    - www.bniplus.co.uk/
- Cumulative Index to Nursing and Allied Health (CINAHL):
    - www.cinahl.com/
- EMCare:
    - www.elsevier.com/wps/find/bibliographicdatabasedescription.cws_home/708272/description#description
- MANTIS (osteopathy and chiropractic):
    - www.healthindex.com/
- OTseeker (systematic reviews and appraised randomized trials in occupational therapy) (free on the internet):
    - www.otseeker.com/
- Physiotherapy Evidence Database (PEDro) (systematic reviews and appraised randomized trials in physiotherapy) (free on the internet):
    - www.pedro.fhs.usyd.edu.au/

Social and community health and welfare

- AgeLine (free on the internet):
    - www.aarp.org/research/ageline/
- Childdata:
    - www.childdata.org.uk/
- CommunityWISE:
    - www.oxmill.com/communitywise/
- Social Care Online (free on the internet):
    - www.scie-socialcareonline.org.uk/
- Social Services Abstracts:
    - www.csa.com/factsheets/ssa-set-c.php

Social science, education, psychology and psychiatry

- Applied Social Sciences Index and Abstracts (ASSIA):
    - www.csa.com/factsheets/assia-set-c.php
- Campbell Collaboration's Social, Psychological, Educational and Criminological Trials Register (C2-SPECTR) (free on the internet):
    - geb9101.gse.upenn.edu/
- Education Resources Information Center (ERIC) (free on the internet)
    - www.eric.ed.gov/

- PsycINFO:
    - www.apa.org/psycinfo/
- Social Policy and Practice (evidence-based social science research):
    - www.ovid.com/site/catalog/DataBase/1859.pdf
- Sociological Abstracts:
    - www.csa.com/factsheets/socioabs-set-c.php

6.2.1.6 Citation indexes

Science Citation Index / Science Citation Index Expanded is a database that lists published articles from approximately 6,000 major scientific, technical and medical journals and links them to the articles in which they have been cited (a feature known as cited reference searching). It is available online as SciSearch and on the internet as Web of Science. Web of Science is also incorporated in Web of Knowledge. It can be searched as a source database like MEDLINE. It can also be used to identify studies for a review by identifying a known relevant source article, and checking each of the articles citing the source article, to see if they are also relevant to the review. It is a way of searching forward in time from the publication of an important relevant article to identify additional relevant articles published since then. Records also include the listed references from the original record, which in turn are another possible source of relevant trial reports. Citation searching is an important adjunct to database searching and handsearching (Greenhalgh 2005). Information about these products is available at:

- scientific.thomson.com/products/sci/
- scientific.thomson.com/products/wos/
- isiwebofknowledge.com/

A similar database exists for the social sciences known as Social Sciences Citation Index:

- scientific.thomson.com/products/ssci/

In 2004, Elsevier launched an abstract and citation database – Scopus. Scopus covers 15,000 journals (of which over 1,200 are open access journals) and 500 conference proceedings. It contains over 33 million abstracts, and results from nearly 400 million scientific web pages:

- info.scopus.com/overview/what/

6.2.1.7 Dissertations and theses databases

Dissertations and theses are not normally indexed in general bibliographic databases such as MEDLINE or EMBASE but there are exceptions, such as CINAHL, which indexes nursing dissertations. To identify relevant studies published in dissertations or theses it is advisable to search specific dissertation sources: see Box 6.2.c.

Box 6.2.c: Examples of dissertations and theses databases

- ProQuest Dissertations & Theses Database: indexes more than 2 million doctoral dissertations and masters' theses:
    - www.proquest.co.uk/products_pq/descriptions/pqdt.shtml
- Index to Theses in Great Britain and Ireland: lists over 500,000 theses:
    - www.theses.com/

- DissOnline: indexes 50,000 German dissertations:
    - www.dissonline.de/

6.2.1.8 Grey literature databases

There are many definitions of grey literature, but it is usually understood to mean literature that is not formally published in sources such as books or journal articles. Conference abstracts and other grey literature have been shown to be sources of approximately 10% of the studies referenced in Cochrane reviews (Mallett 2002). In a recently updated Cochrane methodology review, all five studies reviewed showed that published trials showed an overall greater treatment effect than grey literature trials (Hopewell 2007b). Thus, failure to identify trials reported in conference proceedings and other grey literature might affect the results of a systematic review.

Conference abstracts are a particularly important source of grey literature and are covered in Section 6.2.2.4.

EAGLE (the European Association for Grey Literature Exploitation), has closed the SIGLE (System for Information on Grey Literature) database, which was one of the most widely-used databases of grey literature. INIST in France (Institute for Scientific and Technical) has launched OpenSIGLE, which provides access to all the former SIGLE records, new data added by EAGLE members and information from Greynet.

- opensigle.inist.fr

The Healthcare Management Information Consortium (HMIC) database contains records from the Library & Information Services department of the Department of Health (DH) in England and the King's Fund Information & Library Service. It includes all DH publications including circulars and press releases. The King's Fund is an independent health charity that works to develop and improve management of health and social care services. The database is considered to be a good source of grey literature on topics such as health and community care management, organizational development, inequalities in health, user involvement, and race and health.

- www.ovid.com/site/catalog/DataBase/99.jsp?top=2&mid=3&bottom=7&subsection=10

The National Technical Information Service (NTIS) provides access to the results of both US and non-US government-sponsored research and can provide the full text of the technical report for most of the results retrieved. NTIS from 1964 is free on the internet.

- www.ntis.gov/

PsycEXTRA is a companion database to PsycINFO in psychology, behavioural science and health. It includes references from newsletters, magazines, newspapers, technical and annual reports, government reports and consumer brochures. PsycEXTRA is different from PsycINFO in its format, because it includes abstracts and citations plus full text for a major portion of the records. There is no coverage overlap with PsycINFO.

- www.apa.org/psycextra/

6.2.2 Journals and other non-bibliographic database sources

6.2.2.1 Handsearching

Handsearching involves a manual page-by-page examination of the entire contents of a journal issue or conference proceedings to identify all eligible reports of trials. In journals, reports of trials may appear in articles, abstracts, news columns, editorials, letters or other text. Handsearching healthcare journals and conference proceedings can be a useful adjunct to searching electronic databases for at least two reasons: 1) not all trial reports are included in electronic bibliographic databases, and 2) even when they are included, they may not contain relevant search terms in the titles or abstracts or be indexed with terms that allow them to be easily identified as trials (Dickersin 1994). Each journal year or conference proceeding should be handsearched thoroughly and competently by a well-trained handsearcher for all reports of trials, irrespective of topic, so that once it has been handsearched it will not need to be searched again. A Cochrane Methodology Review has found that a combination of handsearching and electronic searching is necessary for full identification of relevant reports published in journals, even for those that are indexed in MEDLINE (Hopewell 2007a). This was especially the case for articles published before 1991 when there was no indexing term for randomized trials in MEDLINE and for those articles that are in parts of journals (such as supplements and conference abstracts) which are not routinely indexed in databases such as MEDLINE.

To facilitate the identification of all published trials The Cochrane Collaboration has organized extensive handsearching efforts, predominantly through CRGs, Fields and Cochrane Centres. The US Cochrane Center oversees prospective registration of all potential handsearching and maintains files of handsearching activity in the Master List (Journals) and the Master List (Conference Proceedings) (see apps1.jhsph.edu/cochrane/masterlist.asp). Over 3,000 journals have been, or are being, searched within the Collaboration. The Master Lists enable search progress to be recorded and monitored for each title and also prevent duplication of effort which might occur if the same journal or conference proceeding were to be searched by more than one group or individual.

Cochrane entities and authors can prioritize handsearching based on where they expect to identify the most trial reports. This prioritization can be informed by searching CENTRAL, MEDLINE and EMBASE in a topic area and identifying which journals appear to be associated with the most retrieved citations. Preliminary evidence suggests that most of the journals with a high yield of trial reports are indexed in MEDLINE (Dickersin 2002) but this may reflect the fact that Cochrane contributors have concentrated early efforts on searching these journals. Therefore, journals not indexed in MEDLINE or EMBASE should also be considered for handsearching.

Authors are not routinely expected to handsearch journals for their reviews but they should discuss with their Trials Search Co-ordinator whether in their particular case handsearching of any journals or conference proceedings might be beneficial. Authors who wish to handsearch journals or conference proceedings should consult their Trials Search Co-ordinator who can determine whether the journal or conference proceedings has already been searched, and, if it has not, they can register the search on the relevant Master List and provide training in handsearching. Training material is available on the US Cochrane Center web site (apps1.jhsph.edu/cochrane/handsearcher_res.htm).

All correspondence regarding the initiation, progress and status of a journal or conference proceeding search should be between the CRG Trials Search Co-ordinator and staff at the US Cochrane Center.

6.2.2.2 Full text journals available electronically

The full text of an increasing number of journals is available electronically on a subscription basis or free of charge on the internet. In addition to providing a convenient method for retrieving the full article of already identified records, full-text journals can also be searched electronically, depending

on the search interface, in a similar way to the way database records can be searched in a bibliographic database.

It is important to specify if the full text of a journal has been searched electronically. Some journals omit sections of the print version, for example letters, from the electronic version and some include extra articles in electronic format only.

Most academic institutions subscribe to a wide range of electronic journals and these are therefore available free of charge at the point of use to members of those institutions. Review authors should seek advice about electronic journal access from the library service at their local institution. Some professional organizations provide access to a range of journals as part of their membership package. In some countries similar arrangements exist for health service employees through national licences. There are also a number of international initiatives to provide free or low-cost online access to full-text journals (and databases) over the internet, including the Health InterNetwork Access to Research Initiative (HINARI), the International Network for the Availability of Scientific Publications (INASP) and Electronic Information for Libraries (eIFL). For further information on these initiatives see Section 6.2.1.1.

Examples of some full-text journal sources that are available worldwide free of charge without subscription are given in Box 6.2.d.

It is recommended that a local electronic copy or print copy be taken and filed of any possibly relevant article found electronically for subscription journals, as the subscription to that journal may not be in perpetuity. The journal may cease publication or change publishers and access to previously available articles may cease. The same applies to journals available free of charge on the internet, as the circumstances around availability of specific journals might change.

Box 6.2.d: Examples of full-text journal sources available worldwide without charge

- BioMed Central:
    - www.biomedcentral.com/browse/journals/
- Public Library of Science (PLoS):
    - www.plos.org/journals/
- PubMed Central (PMC):
    - www.pubmedcentral.nih.gov/

Web sites listing journals offering free full-text access include:

- Free Medical Journals:
    - freemedicaljournals.com/
- HighWire Press:
    - highwire.stanford.edu/lists/freeart.dtl

6.2.2.3 Tables of contents

Many journals, even those that are available by subscription only, offer Table of Contents (TOC) services free of charge, normally through e-mail alerts or RSS feeds. In addition a number of organizations offer TOC services: see Box 6.2.e.

Box 6.2.e: Examples of organizations offering Table of Contents (TOC) services

- British Library Direct (free):
  - direct.bl.uk/bld/Home.do
- British Library Direct Plus (subscription):
  - www.bl.uk/reshelp/atyourdesk/docsupply/productsservices/bldplus/
- British Library Inside (to be replaced by British Library Direct Plus) (subscription):
  - www.bl.uk/inside
- Current Contents Conntect (subscription):
  - scientific.thomson.com/products/ccc/
- Scientific Electronic Library Online (SciELO) – Brazil (free):
  - www.scielo.br/
- Zetoc (Z39.50 Table Of Contents) (free as specified below):
  Zetoc provides access to the British Library's Electronic Table of Contents. It is free of charge for members of the Joint Information Systems Committee (JISC)-sponsored higher and further education institutions in the UK and all of NHS Scotland and Northern Ireland:
  - zetoc.mimas.ac.uk/

6.2.2.4 Conference abstracts or proceedings

Although conference proceedings are not indexed in MEDLINE and a number of major databases, they are indexed in the BIOSIS databases (http://www.biosis.org/). Over one-half of trials reported in conference abstracts never reach full publication, and those that are eventually published in full have been shown to be systematically different from those that are never published in full (Scherer 2007). It is, therefore, important to try to identify possibly relevant studies reported in conference abstracts through specialist database sources and by handsearching or electronically searching those abstracts that are made available in print form, on CD-ROM or on the internet. Many conference proceedings are published as journal supplements. Specialist conference abstract sources are listed in Box 6.2.f.

Many conference abstracts are published free of charge on the internet, such as those of the American Society of Clinical Oncology (ASCO):

- www.asco.org/ASCO/Meetings

Box 6.2.f: Examples of specialist conference abstract sources

- Biological Abstracts/RRM (Reports, Reviews, Meetings):
  - scientific.thomson.com/products/barrm/
- British Library Inside (to be replaced by British Library Direct Plus):
  - www.bl.uk/inside
- British Library Direct Plus:
  - www.bl.uk/reshelp/atyourdesk/docsupply/productsservices/bldplus
- ISI Proceedings:
  - scientific.thomson.com/products/proceedings/

6.2.2.5 Other reviews, guidelines and reference lists as sources of studies

Some of the most convenient and obvious sources of references to potentially relevant studies are existing reviews. Copies of previously published reviews on, or relevant to, the topic of interest should be obtained and checked for references to the included (and excluded) studies. As well as the *Cochrane Database of Systematic Reviews* (*CDSR)*, *The Cochrane Library* includes *The Database of Abstracts of Reviews of Effects* (DARE) and the *Health Technology Assessment Database* (HTA Database), both produced by the Centre for Reviews and Dissemination (CRD) at the University of York in the UK. Both databases provide information on published reviews of the effects of health care. As well as being published and updated quarterly in *The Cochrane Library,* more up-to-date versions of these databases are available free of charge on the CRD web site, where they are updated more frequently. For example, for the issue of *The Cochrane Library* published in January 2007, the DARE and HTA records were supplied by CRD staff in November 2006. The January 2007 publication of *The Cochrane Library* was the current issue until April 2007, so the DARE and HTA records in *The Cochrane Library* range between being two months to five months out of date.

- www.crd.york.ac.uk/crdweb

CRD used to produce the CRD Ongoing Reviews Database which was searchable through the UK National Research Register (NRR) but since that was archived in September 2007, records of ongoing reviews have been transferred to the HTA Database.

Reviews and guidelines may also provide useful information about the search strategies used in their development: see Box 6.2.g. Specific evidence-based search services such as Turning Research into Practice (TRIP) can be used to identify reviews and guidelines. For the range of systematic review sources searched by TRIP see:

- www.tripdatabase.com/Aboutus/Publications/index.html?catid=11
- www.guideline.gov

MEDLINE, EMBASE and other bibliographic databases can also be used to identify review articles and guidelines. In MEDLINE, the most appropriate review articles should be indexed under the Publication Type term ‘Meta-analysis’, which was introduced in 1993, or ‘Review’, which was introduced in 1966. Guidelines should be indexed under the Publication Type term ‘Practice Guideline’, which was introduced in 1991. EMBASE also has a thesaurus term ‘Systematic Review’, which was introduced in 2003, and ‘Practice Guideline’, which was introduced in 1994.

There is a so-called ‘Systematic Review’ search strategy or filter on PubMed under the Clinical Queries link:

- www.ncbi.nlm.nih.gov/entrez/query/static/clinical.shtml

It is very broad in its scope and retrieves many references that are not systematic reviews. The strategy is described as follows: “This strategy is intended to retrieve citations identified as systematic reviews, meta-analyses, reviews of clinical trials, evidence-based medicine, consensus development conferences, guidelines, and citations to articles from journals specializing in review studies of value to clinicians.”

- www.nlm.nih.gov/bsd/pubmed_subsets/sysreviews_strategy.html

Search strategies or filters have been developed to identify systematic reviews in MEDLINE (White 2001, Montori 2005) and EMBASE (Wilczynski 2007). Search strategies for identifying systematic reviews in other databases and for identifying guidelines are listed on the InterTASC Information Specialists' Subgroup Search Filter Resource web site.

- www.york.ac.uk/inst/crd/intertasc/sr.htm

As well as searching the references cited in existing systematic reviews and meta-analyses, reference lists of identified studies may also be searched for additional studies (Greenhalgh 2005). Since investigators may selectively cite studies with positive results, reference lists should be used with caution as an adjunct to other search methods (see Chapter 10, Section 10.2.2.3).

Box 6.2.g: Examples of evidence-based guidelines

- Australian National Health and Medical Research Council: Clinical Practice Guidelines:
  - nhmrc.gov.au/publications/subjects/clinical.htm
- Canadian Medical Association – Infobase: Clinical Practice Guidelines:
  - mdm.ca/cpgsnew/cpgs/index.asp
- National Guideline Clearinghouse (US):
  - www.guideline.gov/
- National Library of Guidelines (UK):
  - www.library.nhs.uk/guidelinesFinder/
- New Zealand Guidelines Group:
  - www.nzgg.org.nz
- NICE Clinical Guidelines (UK):
  - www.nice.org.uk/aboutnice/whatwedo/aboutclinicalguidelines/about_clinical_guidelines.jsp

6.2.2.6 Web searching

There is little empirical evidence as to the value of using general internet search engines such as Google to identify potential studies (Eysenbach 2001). Searching research funders' and device manufacturers' web sites might be fruitful. Searching pharmaceutical industry web sites may be useful, in particular their trials registers, covered in Section 6.2.3.3. If internet searches are conducted, it is recommended that review authors should file a print copy or save locally an electronic copy of details of information about any possibly relevant study found on the internet, rather than simply 'book-marking' the site, in case the record of the trial is removed or altered at a later stage. It is important to keep a record of the date the web site was accessed for citation purposes.

6.2.3 Unpublished and ongoing studies

Some completed studies are never published. An association between 'significant' results and publication has been documented across a number of studies, as summarized in Chapter 10 (Section 10.2). Finding out about unpublished studies, and including them in a systematic review when eligible and appropriate, is important for minimizing bias. There is no easy and reliable way to obtain information about studies that have been completed but never published. This situation is improving as a result of a number of initiatives:

- The International Standard Randomised Controlled Trial Number Register scheme launched as the first online service that provided unique numbers to randomized controlled trials in all areas of health care and from all countries around the world and subsequently ClinicalTrials.gov (see Section 6.2.3.1);

- The increasing acceptance on behalf of investigators of the importance of registering trials at inception;
- The support of registration at inception by the leading medical journal publishers and their refusal to subsequently publish reports of trials not properly registered (De Angelis 2004, De Angelis 2005);
- The US National Institutes for Health (NIH) Public Access Policy (see publicaccess.nih.gov/), which until December 2007 was voluntary but now requires that "all investigators funded by the NIH submit or have submitted for them to the National Library of Medicine's PubMed Central an electronic version of their final, peer-reviewed manuscripts upon acceptance for publication to be made publicly available no later than 12 months after the official date of publication".
    - publicaccess.nih.gov/policy.htm

Colleagues can be an important source of information about unpublished studies, and informal channels of communication can sometimes be the only means of identifying unpublished data. Formal letters of request for information can also be used to identify completed but unpublished studies. One way of doing this is to send a comprehensive list of relevant articles along with the inclusion criteria for the review to the first author of reports of included studies, asking if they know of any additional studies (published or unpublished) that might be relevant. It may also be desirable to send the same letter to other experts and pharmaceutical companies or others with an interest in the area. It should be borne in mind that asking researchers for information about completed but never published studies has not always been found to be fruitful (Hetherington 1989, Horton 1997) though some researchers have reported that this is an important method for retrieving studies for systematic reviews (Royle 2003, Greenhalgh 2005). Some organizations set up web sites for systematic review projects listing the studies identified to date and inviting submission of information on studies not already listed. It has also been suggested that legislation such as the Freedom of Information Acts in countries such as the UK and the US might be used to gain access to information about unpublished trials (Bennett 2003, MacLean 2003).

It is also important to identify ongoing studies, so that when a review is later updated these can be assessed for possible inclusion. Information about possibly relevant ongoing studies should be included in the review in the 'Characteristics of ongoing studies' table (see Chapter 4, Section 4.6.5). Awareness of the existence of a possibly relevant ongoing study might also affect decisions with respect to when to update a specific review. Unfortunately, no single, comprehensive, centralized register of ongoing trials exists (Manheimer 2002). Efforts have, however, been made by a number of organizations, including organizations representing the pharmaceutical industry and pharmaceutical companies themselves, to begin to provide central access to ongoing trials and in some cases trial results on completion, either on a national or international basis. In an effort to improve this situation, the World Health Organization (WHO) launched the International Clinical Trials Registry Platform Search Portal in May 2007 to search across a range of trials registers, similar to the initiative launched some years earlier by Current Controlled Trials with their so-called *meta*Register. Currently (as at June 2008) the WHO portal only searches across three primary registers (the Australian and New Zealand Clinical Trials Registry, ClinicalTrials.gov and the Current Controlled Trials International Standard Randomised Controlled Trial Number Register) but it is anticipated that other registers will be included as the project progresses.

6.2.3.1 National and international trials registers

Box 6.2.h lists national and international trials registers.

In addition, Drugs@FDA provides information about most of the drugs approved in the US since 1939. For those approved more recently (from 1998), there is often a 'review', which contains the scientific analyses that provided the basis for approval of the new drug.

- www.accessdata.fda.gov/scripts/cder/drugsatfda/index.cfm

Other national and regional drug approval agencies may also be useful sources of trial information.

Box 6.2.h: Examples of national and international trials registers

- The Association of the British Pharmaceutical Industry (ABPI) – Pharmaceutical Industry Clinical Trials database:
  - www.cmrinteract.com/clintrial/
- The Australian New Zealand Clinical Trials Registry:
  - www.anzctr.org.au/
- CenterWatch Clinical Trials Listing Service:
  - www.centerwatch.com/
- Chinese Clinical Trial Register:
  - www.chictr.org/Default.aspx
- ClinicalTrials.gov register:
  - clinicaltrials.gov/
- Community Research & Development Information Service (of the European Union) (trials and other research):
  - cordis.europa.eu/en/home.html
- Current Controlled Trials *meta*Register of Controlled Trials (*m*RCT) – active registers:
  - www.controlled-trials.com/mrct/
- Current Controlled Trials *meta*Register of Controlled Trials (*m*RCT) – archived registers:
  - www.controlled-trials.com/mrct/archived
- European Medicines Agency (EMEA):
  - www.emea.europa.eu/index/indexh1.htm
- German trials register – not yet launched. Final agreement reached 30 August 2007 – will be included under the WHO International Clinical Trials Registry Platform Search Portal – for further details as and when available see:
  - www.who.int/trialsearch
- Hong Kong clinical trials register – HKClinicalTrials.com:
  - www.hkclinicaltrials.com/
- Indian clinical trials registry – Clinical Trials Registry – India (CTRI):
  - www.ctri.in
- International Clinical Trials Registry Platform Search Portal:
  - www.who.int/trialsearch
- International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) Clinical Trials Portal:
  - www.ifpma.org/clinicaltrials.html
- International Standard Randomised Controlled Trial Number Register:

  - www.controlled-trials.com/isrctn/
- Netherlands trial register (Nederlands Trialregister – in Dutch):
  - www.trialregister.nl/trialreg/index.asp
- South African National Clinical Trial Register:
  - www.sanctr.gov.za/
- UK Clinical Research Network Portfolio Database:
  - portal.nihr.ac.uk/Pages/Portfolio.aspx
- UK Clinical Trials Gateway:
  - www.controlled-trials.com/ukctr/
- UK National Research Register (NRR) (trials and other research – archived September 2007 – see UK Clinical Trials Gateway):
  - portal.nihr.ac.uk/Pages/NRRArchive.aspx
- University hospital Medical Information Network (UMIN) Clinical Trials Registry (for Japan) – UMIN CTR:
  - www.umin.ac.jp/ctr/

6.2.3.2 Subject-specific trials registers

There are many condition-specific trials registers, especially in the field of cancer – which are too numerous to list. They can be identified by searching the internet and by searching within some of the resources listed above such as the Current Controlled Trials *meta*Register of Controlled Trials (*m*RCT).

6.2.3.3 Pharmaceutical industry trials registers

Some pharmaceutical companies make available information about their clinical trials though their own web sites, either instead of or in addition to the information they make available through national or international web sites such as those listed above. Some examples are included in Box 6.2.i.

Box 6.2.i: Examples of pharmaceutical industry trials registers

- AstraZeneca Clinical Trials web site:
  - www.astrazenecaclinicaltrials.com/
- Bristol-Myers Squibb Clinical Trial Registry:
  - ctr.bms.com/ctd/registry.do
- Eli Lilly and Company Clinical Trial Registry (also includes trial results)
  - www.lillytrials.com/
- GlaxoSmithKline clinical trial register:
  - ctr.gsk.co.uk/medicinelist.asp
- NovartisClinicalTrials.com:
  - www.novartisclinicaltrials.com/webapp/etrials/home.do
- Roche Clinical Trial Protocol Registry:
  - www.roche-trials.com/registry.html

- Wyeth Clinical Trial Listings:
    - www.wyeth.com/ClinicalTrialListings

6.2.3.4 Trials results registers and other sources

Registers of the results of completed trials are a more recent phenomenon, following on from ongoing trials registers that simply list details of the trial. They are of particular value because trial results are not always published, and even if published are not always published in full. Recent legislation in the US known as Section 801 of the Food and Drug Administration Amendments Act of 2007 (FDAAA 801), enacted in September 2007, called for expanding ClinicalTrials.gov and adding a clinical trial results database. Examples of trials results registers are provided in Box 6.2.j.

In addition, Clinical Trial Results is a web site that hosts slide presentations from clinical trialists reporting the results of clinical trials:

- www.clinicaltrialresults.org/

Box 6.2.j: Examples of trials results registers

- International Federation of Pharmaceutical Manufacturers and Associations (IFPMA) Clinical Trials Portal:
    - www.ifpma.org/clinicaltrials.html
- PhRMA Clinical Study Results Database:
    - www.clinicalstudyresults.org/about
- Bristol-Myers Squibb Clinical Trial Results:
    - ctr.bms.com/ctd/results.do
- Eli Lilly and Company Clinical Trial Registry:
    - www.lillytrials.com/
- Roche Clinical Trials Results Database:
    - www.roche-trials.com/results.html
- Wyeth Clinical Trial Results:
    - www.wyeth.com/ClinicalTrialResults

6.2.4 Summary points

- Cochrane review authors should seek advice from their Trials Search Co-ordinator on sources to search.
- CENTRAL is considered to be the best single source of reports of trials for inclusion in Cochrane reviews.
- The three bibliographic databases generally considered to be the most important sources to search for studies for inclusion in Cochrane reviews are CENTRAL, MEDLINE and EMBASE.
- National, regional and subject-specific databases should be selected for searching according to the topic of the review.
- Conference abstracts and other grey literature can be an important source of studies for inclusion in reviews.

- Reference lists in other reviews, guidelines, included (and excluded) studies and other related articles should be searched for additional studies.
- Efforts should be made to identify unpublished studies.
- Ongoing trials should be identified and tracked for possible inclusion in reviews on completion.
- Trials registers and trials results registers are an important source of ongoing and unregistered trials.

6.3 Planning the search process

6.3.1 Involving Trials Search Co-ordinators and healthcare librarians in the search process

It is the responsibility of each CRG to support review authors in identifying reports of studies for inclusion in their reviews, and most CRGs employ a Trials Search Co-ordinator to fulfil this role (see Section 6.1.1.1). Most CRGs offer support to authors in study identification from the early planning stage to the final write-up of the review for publication in the *CDSR*. This support might include designing search strategies or advising on their design, running searches, in particular in databases not available to the review author at their institution, and providing review authors with lists of references to studies from the CRG's Specialized Register and possibly from other databases. The range of services offered varies across CRGs according to the resources available. Review authors are, therefore, encouraged to contact the Trials Search Co-ordinator of their CRG at the earliest stage for advice and support.

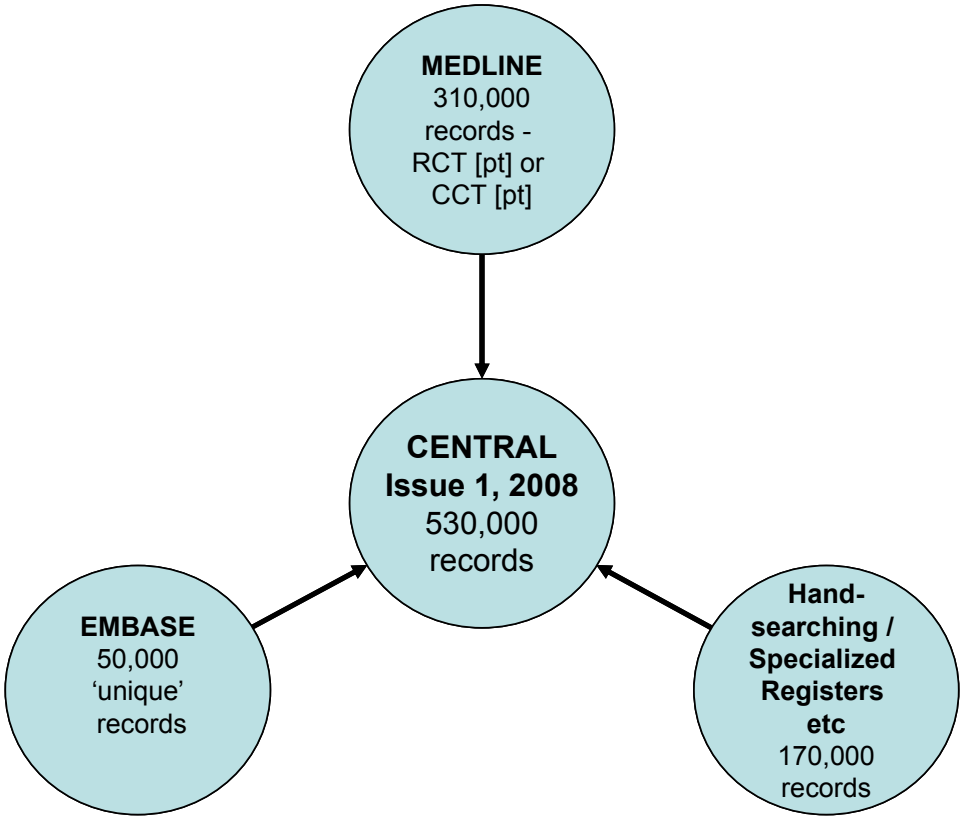
If authors are conducting their own searches, they should seek advice from their Trials Search Co-ordinator with respect to which database(s) to search and the exact strategies to be run. It should also be borne in mind that the search process needs to be documented in enough detail throughout to ensure that it can be reported correctly in the review, to the extent that all the searches of all the databases are reproducible. The full search strategies for each database should be included in the review in an Appendix. It is, therefore, important that review authors should save all search strategies and take notes at the time to enable the completion of that section at the appropriate time. For further guidance on this, authors should contact their Trials Search Co-ordinator, and see Section 6.6.

If the CRG is currently without a Trials Search Co-ordinator it is recommended that review authors seek guidance from a healthcare librarian or information specialist, where possible with experience of supporting systematic reviews.

6.3.2 Collaboration-wide search initiatives

In planning the search process it is necessary to take into account what other searching has already been undertaken to avoid unnecessary duplication of effort. For example, considerable efforts over the years have gone into searching MEDLINE and EMBASE and incorporating reports of trials from these two major international databases into the Cochrane Central Register of Controlled Trials (CENTRAL). It is necessary, therefore, that any additional searching for a specific review should take into account what has gone before. Figure 6.3.a illustrates the contents of CENTRAL.

Figure 6.3.a: Illustration of the contents of CENTRAL



6.3.2.1 What is in The Cochrane Central Register of Controlled Trials (CENTRAL) from MEDLINE?

CENTRAL contains all records from MEDLINE indexed with the Publication Type term 'Randomized Controlled Trial' or 'Controlled Clinical Trial' that are indexed as human studies. These records are downloaded quarterly from MEDLINE by Wiley-Blackwell as part of the build of CENTRAL for publication in *The Cochrane Library*. For further details see:

- www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/CENTRALHelpFile.html

A substantial proportion of the MEDLINE records coded 'Randomized Controlled Trial' or 'Controlled Clinical Trial' in the Publication Type field have been coded as a result of the work of The Cochrane Collaboration (Dickersin 2002). Handsearch results from Cochrane entities, for journals indexed in MEDLINE, have been sent to the US National Library of Medicine (NLM), where the MEDLINE records have been re-tagged with the publication types 'Randomized Controlled Trial' or 'Controlled Clinical Trial' as appropriate. In addition, the US Cochrane Center (formerly the New England Cochrane Center, Providence Office and the Baltimore Cochrane Center) and the UK Cochrane Centre have conducted an electronic search of MEDLINE from 1966–2004 to identify reports of randomized controlled trials, identifiable from the MEDLINE titles and/or abstracts, not already indexed as such, using the first two phases of the Cochrane Highly Sensitive Search Strategy first published in 1994 (Dickersin 1994) and subsequently updated and included in the *Handbook*. The free text terms used were: clinical trial; (singl$ OR doubl$ OR trebl$ OR tripl$) AND (mask$ OR blind$); placebo$; random$. The $ sign indicates the use of a truncation symbol. The following subject index terms (MeSH) used were exploded: randomized controlled trials; random allocation; double-blind method; single-blind method; clinical trials; placebos. The following subject heading

(MeSH) was used unexploded: research design. The Publication Type terms used were: randomized controlled trial; controlled clinical trial; clinical trial.

A test was carried out using the terms in phase three of the 1994 Cochrane Highly Sensitive Search Strategy but the precision of those terms, having already searched on all the terms in phases one and two as listed above, was considered to be too low to warrant using these terms for the above project (Lefebvre 2001). It was, however, recognized that some of these terms might be useful when combined with subject terms to identify studies for some specific reviews (Eisinga 2007).

The above search was limited to humans. The following years were completed by the US Cochrane Center (1966–1984; 1998–2004) and by the UK Cochrane Centre (1985–1997). The results have been forwarded to the NLM and re-tagged in MEDLINE and are thus included in CENTRAL. This project is currently on hold. If the US Cochrane Center can attract funding for this project they will continue the electronic search of records entered into MEDLINE in 2005 and beyond. Any updates to this situation will be described in the CENTRAL Creation Details file in *The Cochrane Library*:

- www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/CENTRALHelpFile.html

CENTRAL includes from MEDLINE not only reports of trials that meet the more restrictive Cochrane definition for a controlled clinical trial (Box 6.3.a) but also trial reports that meet the less restrictive original NLM definition (Box 6.3.b), which used to include historical comparisons. There is currently no method of distinguishing, either in CENTRAL or in MEDLINE, which of these records meet the more restrictive Cochrane definition, as they are all indexed with the Publication Type term 'Controlled Clinical Trial'.

Box 6.3.a: Cochrane definitions and criteria for randomized controlled trials (RCTs) and controlled clinical trials (CCTs)

Records identified for inclusion should meet the eligibility criteria devised and agreed in November 1992, which were first published, in 1994, in the first version of the *Handbook* (see Chapter 1, Section 1.4). According to these eligibility criteria:

A trial is eligible if, on the basis of the best available information (usually from one or more published reports), it is judged that:

- the individuals (or other units) followed in the trial were definitely or possibly assigned prospectively to one of two (or more) alternative forms of health care using
  - random allocation or
  - some quasi-random method of allocation (such as alternation, date of birth, or case record number).

Trials eligible for inclusion are classified according to the reader's degree of certainty that random allocation was used to form the comparison groups in the trial. If the author(s) state explicitly (usually by some variant of the term 'random' to describe the allocation procedure used) that the groups compared in the trial were established by random allocation, then the trial is classified as a RCT (randomized controlled trial). If the author(s) do not state explicitly that the trial was randomized, but randomization cannot be ruled out, the report is classified as a CCT (controlled clinical trial). The classification CCT is also applied to quasi-randomized studies, where the method of allocation is known but is not considered strictly random, and possibly quasi-randomized trials. Examples of quasi-random methods of assignment include alternation, date of birth, and medical record number.

The classification as RCT or CCT is based solely on what the author has written, not on the reader's

interpretation; thus, it is not meant to reflect an assessment of the true nature or quality of the allocation procedure. For example, although 'double-blind' trials are nearly always randomized, many trial reports fail to mention random allocation explicitly and should therefore be classified as CCT.

Relevant reports are reports published in any year, of studies comparing at least two forms of health care (healthcare treatment, healthcare education, diagnostic tests or techniques, a preventive intervention, etc.) where the study is on either living humans or parts of their body or human parts that will be replaced in living humans (e.g., donor kidneys). Studies on cadavers, extracted teeth, cell lines, etc. are not relevant. *Searchers should identify all controlled trials meeting these criteria regardless of relevance to the entity with which they are affiliated.*

The highest possible proportion of all reports of controlled trials of health care should be included in CENTRAL. Thus, those searching the literature to identify trials should give reports the benefit of any doubts. Review authors will decide whether to include a particular report in a review.

Box 6.3.b: US National Library of Medicine 2008 definitions for the Publication Type terms 'Randomized Controlled Trial' and 'Controlled Clinical Trial'

Randomized Controlled Trial

Work consisting of a clinical trial that involves at least one test treatment and one control treatment, concurrent enrolment and follow-up of the test- and control-treated groups, and in which the treatments to be administered are selected by a random process, such as the use of a random-numbers table.

Controlled Clinical Trial

Work consisting of a clinical trial involving one or more test treatments, at least one control treatment, specified outcome measures for evaluating the studied intervention, and a bias-free method for assigning patients to the test treatment. The treatment may be drugs, devices, or procedures studied for diagnostic, therapeutic, or prophylactic effectiveness. Control measures include placebos, active medicine, no-treatment, dosage forms and regimens, historical comparisons, etc. When randomization using mathematical techniques, such as the use of a random-numbers table, is employed to assign patients to test or control treatments, the trial is characterized as a 'Randomized Controlled Trial'.

6.3.2.2 What is in The Cochrane Central Register of Controlled Trials (CENTRAL) from EMBASE?

In a study similar to that described above for MEDLINE, a search of EMBASE has been carried out by the UK Cochrane Centre for reports of trials not indexed as trials in MEDLINE (Lefebvre 2008). (Trials indexed as such in MEDLINE are already included in CENTRAL as described in Section 6.3.2.1, and are therefore de-duplicated against the EMBASE records as part of the search process.) The following terms are those currently used for the project and have been searched for the years 1980 to 2006: free-text terms: random$; factorial$; crossover$; cross over$; cross-over$; placebo$; doubl$ adj blind$; singl$ adj blind$; assign$; allocat$; volunteer$; and index terms, known as EMTREE terms: crossover-procedure; double-blind procedure; randomized controlled trial; single-blind procedure. A search for the years 1974 to 1979 inclusive has also been completed for the free-text terms: random$; factorial$; crossover$ and placebo$. The $ sign indicates the use of a truncation symbol.

These searches have yielded a total of 80,000 reports of trials not, at the time of the search, indexed as reports of trials in MEDLINE. All of these records are now published in CENTRAL, under contract between Elsevier, the publishers of EMBASE, and The Cochrane Collaboration. Of these 80,000 records, 50,000 are 'unique' to CENTRAL, that is they are not already included in CENTRAL with the records sourced from MEDLINE. This search is updated annually. Updates are described in the CENTRAL Creation Details file in *The Cochrane Library*:

- www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/CENTRALHelpFile.html

and the What's New section on *The Cochrane Library* home page:

- www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/HOME

6.3.2.3 What is in The Cochrane Central Register of Controlled Trials (CENTRAL) from other databases and handsearching?

Other general healthcare databases such as those published in Australia and China have undergone similar systematic searches to identify reports of trials for CENTRAL. The Australasian Cochrane Centre co-ordinated a search of the National Library of Australia's Australasian Medical Index from 1966 (McDonald 2002). This search has recently been updated to include records added up to 2007. The Chinese Cochrane Center, with support from the Australasian Cochrane Centre, co-ordinated a search of the Chinese Biomedical Literature Database from 1999 to 2001. In an ongoing project, the Chinese Cochrane Center, with support from the UK Cochrane Centre, is searching a number of Chinese sources with a view to including these records in CENTRAL. Similarly, the Brazilian Cochrane Centre in collaboration with the Regional Library of Medicine in Brazil (BIblioteca REgional de MEdicina – BIREME) is planning to co-ordinate a search of the Pan American Health Organization's database LILACS (Latin American Caribbean Health Sciences Literature).

Each of the Cochrane Centres has the responsibility for searching the general healthcare literature of its country or region. The CRGs and Fields are responsible for co-ordinating searching of the specialist healthcare literature in their areas of interest. More than 3000 journals have been, or are being, handsearched. Identified trial reports that are not relevant to a CRG's scope and thus are not appropriate for their Specialized Register (see below) are forwarded to Wiley-Blackwell as handsearch results. Handsearch records can be identified in CENTRAL as they are assigned the tag HS-HANDSRCH or HS-PRECENTRL.

- www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/CENTRALHelpFile.html

6.3.2.4 What is in The Cochrane Central Register of Controlled Trials (CENTRAL) from Specialized Registers of Cochrane Review Groups and Fields?

It is an 'essential core function' of CRGs that their "editorial bases develop and maintain a Specialized Register, containing all relevant studies in their area of interest, and submit this to CENTRAL on a quarterly basis", as outlined in Section 3.2.1.5 'Core functions of Cochrane Review Groups' in The Cochrane Manual (www.cochrane.org/admin/manual.htm).

The Specialized Register serves to ensure that individual review authors within the CRG have easy and reliable access to trials relevant to their review topic, normally through their Trials Search Co-ordinator. CRGs use the methods described in this Chapter of the *Handbook* to identify trials for their Specialized Registers. Most CRGs also have systems in place to ensure that any additional eligible reports identified by authors for their review(s) are contributed to the CRG's Specialized Register. The registers are, in turn, submitted for inclusion in CENTRAL on a quarterly basis. Thus, records included in the Specialized Register of one CRG become accessible to all other CRGs through CENTRAL. Many Fields also develop subject-specific Specialized Registers and submit them for

inclusion in CENTRAL as described above. To identify records in CENTRAL from within a specific Specialized Register it is possible to search on the Specialized Register tag, such as SR-STROKE. A list of all the Specialized Register tags can be found in the 'Appendix: Review Group or Field/Network Specialized Register Codes' in the 'CENTRAL Creation Details' Help File in *The Cochrane Library*:

- www3.interscience.wiley.com/cgi-bin/mrwhome/106568753/CENTRALHelpFile.html

Records in a CRG's Specialized Register will often contain coding and other information not included in CENTRAL, so the Trials Search Co-ordinator will often be able to identify additional records in their Specialized Register, which could not be identified by searching in CENTRAL, by searching for these codes in the Specialized Register. Conversely, the search functionality of the bibliographic or other software used to manage Specialized Registers is usually less sophisticated than the search functionality available in *The Cochrane Library* so a search of CENTRAL will retrieve records from the Specialized Register that may not be easily retrievable from within the Specialized Register itself. It is therefore recommended that both CENTRAL and the Specialized Register itself are searched separately to maximize retrieval.

6.3.3 Searching CENTRAL, MEDLINE and EMBASE: specific issues

It is recommended that for all Cochrane reviews, CENTRAL and MEDLINE should be searched, as a minimum, together with EMBASE if it is available to either the CRG or the review author.

6.3.3.1 Searching The Cochrane Central Register of Controlled Trials (CENTRAL): specific issues

CENTRAL is comprised of records from a wide range of sources (see Section 6.2.1.2 and 6.3.2 and subsections), so there is no consistency in the format or content of the records.

The 310,000 records sourced from MEDLINE are best retrieved by a combination of Medical Subject Heading (MeSH) and free-text terms. The other records, including the 50,000 records sourced from EMBASE, are best retrieved using free-text searches across all fields.

Most of the records that do not come from MEDLINE or EMBASE (about 170,000 in *The Cochrane Library* Issue 1, 2008) do not have abstracts or any indexing terms. To retrieve these records, which consist predominantly of titles only, it is necessary to carry out a very broad search consisting of a wide range of free-text terms, which may be considered too broad to run across the whole of CENTRAL.

It is possible to identify the records that have been sourced from MEDLINE and EMBASE by searching in CENTRAL for those records that have PubMed or EMBASE accession numbers. It is possible then to exclude these records from a broad search of CENTRAL, as illustrated in the example in Box 6.3.c.

For general information about searching, which is relevant to searching CENTRAL, see Section 6.4.

Box 6.3.c: Example of exclusion of MEDLINE and EMBASE records when searching CENTRAL

Note: the example is for illustrative purposes only. A search of CENTRAL for a systematic review on this topic would require a wide range of alternative terms for both tamoxifen and breast cancer.

#1	"accession number" near pubmed
#2	"accession number" near2 embase
#3	#1 or #2
#4	tamoxifen
#5	(breast near cancer)
#6	#4 and #5
#7	#6 not #3

6.3.3.2 Searching MEDLINE and EMBASE: specific issues

Despite the fact that both MEDLINE and EMBASE have been searched systematically for reports of trials and that these reports of trials have been included in CENTRAL, as described in Sections 6.3.2.1 and 6.3.2.2, supplementary searches of both MEDLINE and EMBASE are recommended. Any such searches, however, should be undertaken in the knowledge of what searching has already been conducted to avoid duplication of effort.

Searching MEDLINE

There is a delay of some months between records being indexed in MEDLINE and appearing indexed as reports of trials in CENTRAL, since CENTRAL is only updated quarterly. For example, for the issue of *The Cochrane Library* published in January 2007, the MEDLINE records were downloaded by Wiley-Blackwell staff in November 2006. The January 2007 publication of *The Cochrane Library* was the current issue until April 2007, so the MEDLINE records range between being two to five months out of date. The most recent months of MEDLINE should, therefore, be searched, at least for records indexed as either 'Randomized Controlled Trial' or 'Controlled Clinical Trial' in the Publication Type, to identify those records recently indexed as RCTs or CCTs in MEDLINE.

Additionally, the most recent year to be searched under the project to identify reports of trials in MEDLINE and send them back to the US National Library of Medicine for re-tagging was 2004, so records added to MEDLINE during and since 2005 should be searched using one of the search strategies described in Section 6.4.11.1.

Finally, for extra sensitivity, or where the use of a randomized trial 'filter' is not appropriate, review authors should search MEDLINE for all years using subject terms only.

It should be remembered that the MEDLINE re-tagging project described in Section 6.3.2.1 assessed whether the records identified were reports of trials on the basis of the title and abstract only, so any supplementary search of MEDLINE that is followed up by accessing the full text of the articles will identify additional reports of trials, most likely through the methods sections, that were not identified through the titles or abstracts alone.

For guidance on running separate search strategies in the MEDLINE-indexed versions of MEDLINE and the versions of MEDLINE containing 'in process' and other non-indexed records please refer to Section 6.4.11.1.

Any reports of trials identified by the review author can be submitted to the Trials Search Co-ordinator who can ensure that they are added to CENTRAL. Any errors, in respect of records indexed as trials in MEDLINE that on the basis of the full article are definitely not reports of trials according to the definitions used by the National Library of Medicine (NLM) (see Section 6.3.2.1), should also be reported to the Trials Search Co-ordinator, so they can be referred to the NLM and corrected.

For general information about searching, which is relevant to searching MEDLINE, see Section 6.4.

Searching EMBASE

The project to identify reports of trials in EMBASE for inclusion in CENTRAL, described in Section 6.3.2.2, is carried out on an annual basis, so there is a time lag of approximately one to two years with respect to EMBASE records appearing in CENTRAL. The last two years of EMBASE should, therefore, be searched to cover work still in progress. Some suggested search terms are listed in Section 6.3.2.2. A search filter designed by the McMaster Hedges Team is also available (Wong 2006).

Finally, for extra sensitivity, or where the use of a randomized trial 'filter' is not appropriate, review authors should search EMBASE for all years using subject terms only, as described under similar circumstances for MEDLINE above. It should be remembered that the EMBASE project described above assessed whether the records identified were reports of trials on the basis of the title and abstract only, in the same way as the MEDLINE project described above. Therefore, any supplementary search of EMBASE that is followed up by accessing the full text of the articles will identify additional reports of trials, most likely through the methods sections, that were not identified through the titles or abstracts alone.

For general information about searching, which is relevant to searching EMBASE, see Section 6.4.

6.3.4 Summary points

- Cochrane review authors should seek advice from their Trials Search Co-ordinator throughout the search process.
- It is recommended that for all Cochrane reviews CENTRAL and MEDLINE should be searched, as a minimum, together with EMBASE if it is available to either the CRG or the review author.
- The full search strategies for each database searched will need to be included in an Appendix of the review, so all search strategies should be saved, and notes taken of the number of records retrieved for each database searched.
- CENTRAL contains over 350,000 records from MEDLINE and EMBASE, so care should be taken when searching MEDLINE and EMBASE to avoid unnecessary duplication of effort.
- MEDLINE should be searched from 2005 onwards inclusive using one of the revised and updated Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE as outlined in Section 6.4.11.1.
- EMBASE should be searched for the most recent two years as outlined in Section 6.4.11.2.
- Additional studies can be identified in MEDLINE and EMBASE by searching across the years already searched for CENTRAL, by obtaining the full article and by reading, in particular, the methods section.

6.4 Designing search strategies

6.4.1 Designing search strategies – an introduction

This section highlights some of the issues to consider when designing search strategies, but does not adequately address the many complexities in this area. It is in particular in this aspect of searching for studies that the skills of a Trials Search Co-ordinator or healthcare librarian are highly recommended. Many of the issues highlighted below relate to both the methodological aspect of the search (such as identifying reports of randomized trials) and the subject of the search. For a search to be robust both aspects require equal attention to be sure that relevant records are not missed.

The eligibility criteria for studies to be included in the review will inform how the search is conducted (see Chapter 5). The eligibility criteria will specify the types of designs, types of participants, types of intervention (experimental and comparator) and, in some cases, the types of outcomes to be addressed. Issues to consider in planning a search include the following:

- whether the review is limited to randomized trials or whether other study designs will be included (see also Chapter 13);
- the requirement to identify adverse effects data (see also Chapter 14);
- the nature of the intervention(s) being assessed;
- any geographic considerations such as the need to search the Chinese literature for studies in Chinese herbal medicine;
- the time period when any evaluations of these interventions may have taken place; and
- whether data from unpublished studies are to be included.

6.4.2 Structure of a search strategy

The structure of a search strategy should be based on the main concepts being examined in a review. For a Cochrane review, the review title should provide these concepts and the eligibility criteria for studies to be included will further assist in the selection of appropriate subject headings and text words for the search strategy.

It is usually unnecessary, and even undesirable, to search on every aspect of the review's clinical question (often referred to as PICO – that is Patient (or Participant or Population), Intervention, Comparison and Outcome). Although a research question may address particular populations, settings or outcomes, these concepts may not be well described in the title or abstract of an article and are often not well indexed with controlled vocabulary terms. They generally, therefore, do not lend themselves well to searching. In general databases, such as MEDLINE, a search strategy to identify studies for a Cochrane review will typically have three sets of terms: 1) terms to search for the health condition of interest, i.e. the population; 2) terms to search for the intervention(s) evaluated; and 3) terms to search for the types of study design to be included (typically a 'filter' for randomized trials). CENTRAL, however, aims to contain only reports with study designs possibly relevant for inclusion in Cochrane reviews, so searches of CENTRAL should not use a trials 'filter'. Filters to identify randomized trials and controlled trials have been developed specifically for MEDLINE and guidance is also given for searching EMBASE: see Section 6.4.11 and sub-sections. For reviews of complex interventions, it may be necessary to adopt a different approach, for example by searching only for the population or the intervention (Khan 2001).

6.4.3 Service providers and search interfaces

Both MEDLINE and EMBASE are offered by a number of service providers, via a range of search interfaces; for example Dialog offers both Dialog and DataStar. In addition the US National Library of

Medicine and Elsevier both offer access to their own versions of MEDLINE and EMBASE respectively: MEDLINE through PubMed, which is available free of charge on the internet, and EMBASE through EMBASE.com which is available on subscription only. Search syntax varies across interfaces. For example, to search for the Publication Type term 'Randomized Controlled Trial' in the various search interfaces it is necessary to enter the term as:

randomized controlled trial.pt. (in Ovid)

randomized controlled trial [pt] (in PubMed)

randomized controlled trial in pt (in SilverPlatter)

Many service providers offer links to full-text versions of articles on other publishers' web sites, such as the PubMed 'Links / LinkOut' feature.

6.4.4 Sensitivity versus precision

Searches for systematic reviews aim to be as extensive as possible in order to ensure that as many as possible of the necessary and relevant studies are included in the review. It is, however, necessary to strike a balance between striving for comprehensiveness and maintaining relevance when developing a search strategy. Increasing the comprehensiveness (or sensitivity) of a search will reduce its precision and will retrieve more non-relevant articles.

Sensitivity is defined as the number of relevant reports identified divided by the total number of relevant reports in existence. Precision is defined as the number of relevant reports identified divided by the total number of reports identified.

Developing a search strategy is an iterative process in which the terms that are used are modified, based on what has already been retrieved. There are diminishing returns for search efforts; after a certain stage, each additional unit of time invested in searching returns fewer references that are relevant to the review. Consequently there comes a point where the rewards of further searching may not be worth the effort required to identify the additional references. The decision as to how much to invest in the search process depends on the question a review addresses, the extent to which the CRG's Specialized Register is developed, and the resources that are available. It should be noted, however, that article abstracts identified through a literature search can be 'scan-read' very quickly to ascertain potential relevance. At a conservatively-estimated reading rate of two abstracts per minute, the results of a database search can be 'scan-read' at the rate of 120 per hour (or approximately 1000 over an 8-hour period), so the high yield and low precision associated with systematic review searching is not as daunting as it might at first appear in comparison with the total time to be invested in the review.

6.4.5 Controlled vocabulary and text words

MEDLINE and EMBASE (and many other databases) can be searched using standardized subject terms assigned by indexers. Standardized subject terms (as part of a controlled vocabulary or thesaurus) are useful because they provide a way of retrieving articles that may use different words to describe the same concept and because they can provide information beyond that which is simply contained in the words of the title and abstract. When searching for studies for a systematic review, however, the extent to which subject terms are applied to references should be viewed with caution. Authors may not describe their methods or objectives well and indexers are not always experts in the subject areas or methodological aspects of the articles that they are indexing. In addition, the available indexing terms might not correspond to the terms the searcher wishes to use.

The controlled vocabulary search terms for MEDLINE (MeSH) and EMBASE (EMTREE) are not identical, and neither is the approach to indexing. For example, the pharmaceutical or pharmacological aspects of an EMBASE record are generally indexed in greater depth than the equivalent MEDLINE record, and in recent years Elsevier has increased the number of index terms assigned to each EMBASE record. Searches of EMBASE may, therefore, retrieve additional articles that were not retrieved by a MEDLINE search, even if the records were present in both databases. Search strategies need to be customized for each database.

One way to begin to identify controlled vocabulary terms for a particular database is to retrieve articles from that database that meet the inclusion criteria for the review, and to note common text words and the subject terms the indexers have applied to the articles, which can then be used for a full search. Having identified a key article, additional relevant articles can be located, for example by using the 'Find Similar' option in Ovid or the 'Related Articles' option in PubMed. Additional controlled vocabulary terms should be identified using the search tools provided with the database, such as the Permuted Index under Search Tools in Ovid and the MeSH Database option in PubMed.

Many database thesauri offer the facility to 'explode' subject terms to include more specific terms automatically in the search. For example, a MEDLINE search using the MeSH term BRAIN INJURIES, if exploded, will automatically search not only for the term BRAIN INJURIES but also for the more specific term SHAKEN BABY SYNDROME. As articles in MEDLINE on the subject of shaken baby syndrome should only be indexed with the more specific term SHAKEN BABY SYNDROME and not also with the more general term BRAIN INJURIES it is important that MeSH terms are 'exploded' wherever appropriate, in order not to miss relevant articles. The same principle applies to EMTREE when searching EMBASE and also to a number of other databases. For further guidance on this topic, review authors should consult their Trials Search Co-ordinator or healthcare librarian.

It is particularly important in MEDLINE to distinguish between Publication Type terms and other related MeSH terms. For example, a report of a randomized trial would be indexed in MEDLINE with the Publication Type term 'Randomized Controlled Trial' whereas an article about randomized controlled trials would be indexed with the MeSH term RANDOMIZED CONTROLLED TRIALS AS TOPIC (note the latter is plural). The same applies to other indexing terms for trials, reviews and meta-analyses.

Review authors should assume that earlier articles are even harder to identify than recent articles. For example, abstracts are not included in MEDLINE for most articles published before 1976 and, therefore, text word searches will only apply to titles. In addition, few MEDLINE indexing terms relating to study design were available before the 1990s, so text word searches are necessary to retrieve older records.

In order to identify as many relevant records as possible searches should comprise a combination of subject terms selected from the controlled vocabulary or thesaurus ('exploded' where appropriate) with a wide range of free-text terms.

6.4.6 Synonyms, related terms, variant spellings, truncation and wildcards

When designing a search strategy, in order to be as comprehensive as possible, it is necessary to include a wide range of free-text terms for each of the concepts selected. For example:

- synonyms: 'pressure sore' OR 'decubitus ulcer', etc;

- related terms: ‘brain’ OR ‘head’, etc; and
- variant spellings: ‘tumour’ OR ‘tumor’.

Service providers offer facilities to capture these variations through truncation and wildcards:

- truncation: random* (for random or randomised or randomized or randomly, etc); and
- wildcard: wom?n (for woman or women).

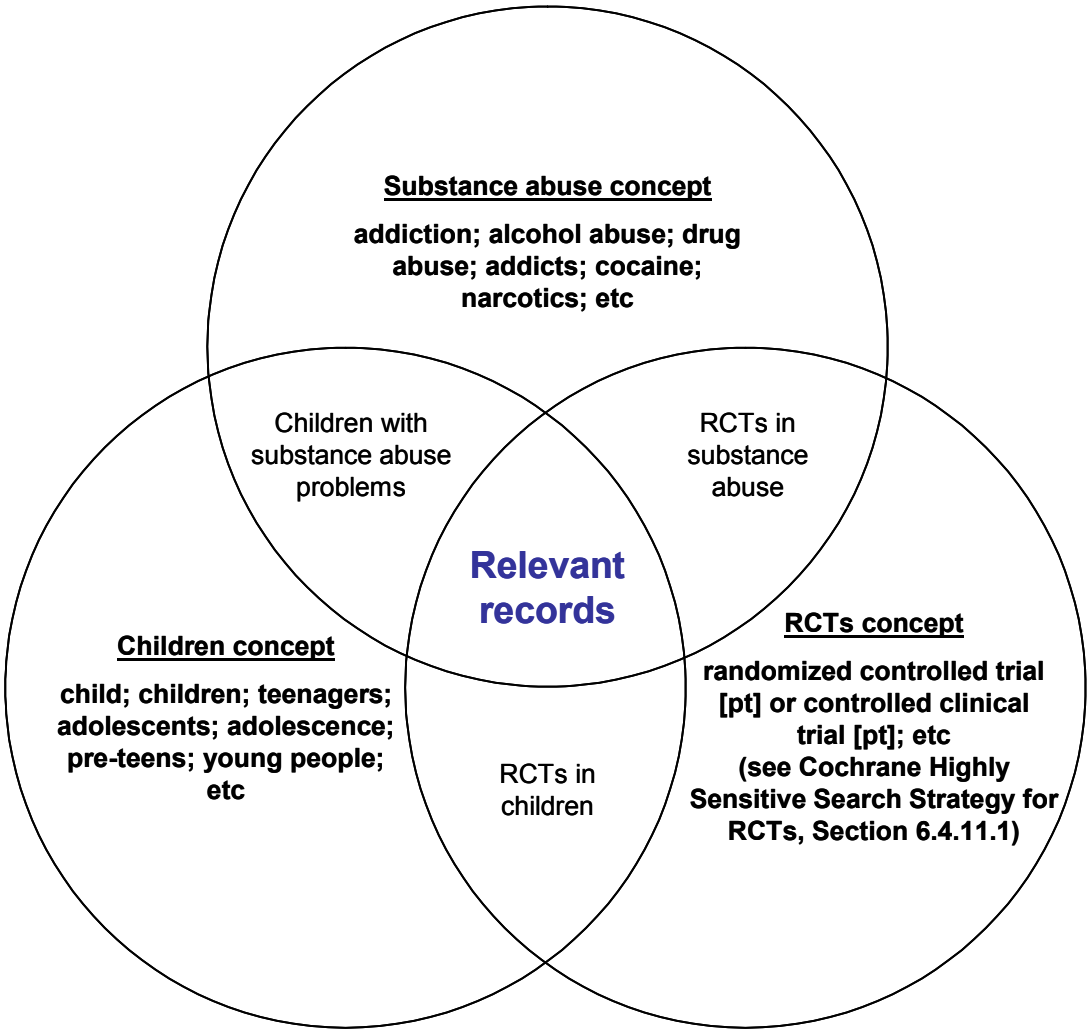
These features vary across service providers. For further details refer to the service provider help files for the database in question.

6.4.7 Boolean operators (AND, OR and NOT)

A search strategy should build up the controlled vocabulary terms, text words, synonyms and related terms for each concept at a time, joining together each of the terms within each concept with the Boolean ‘OR’ operator: see demonstration search strategy Box 6.4.f). This means articles will be retrieved that contain at least one of these search terms. Sets of terms should usually be developed for the healthcare condition, intervention(s) and study design. These three sets of terms can then be joined together with the ‘AND’ operator. This final step of joining the three sets with the ‘AND’ operator limits the retrieved set to articles of the appropriate study design that address both the health condition of interest and the intervention(s) to be evaluated. A note of caution about this approach is warranted however: if an article does not contain at least one term from each of the three sets, it will not be identified. For example, if an index term has not been added to the record for the intervention and the intervention is not mentioned in the title and abstract, the article would be missed. A possible remedy is to omit one of the three sets of terms and decide which records to check on the basis of the number retrieved and the time available to check them. The ‘NOT’ operator should be avoided where possible to avoid the danger of inadvertently removing from the search set records that are relevant. For example, when searching for records indexed as female, ‘NOT male’ would remove any record that was about both males and females.

Searches for Cochrane reviews can be extremely long, often including over 100 search statements. It can be tedious to type in the combinations of these search sets, for example as ‘#1 OR #2 OR #3 OR #4 …. OR #100’. Some service providers offer alternatives to this. For example, in Ovid it is possible to combine sets using the syntax ‘or/1-100’. For those service providers where this is not possible, including *The Cochrane Library* for searches of CENTRAL, it has been recommended that the search string above could be typed in full and saved, for example, as a Word document and the requisite number of combinations copied and pasted into the search as required. Having typed the string with the # symbols as above, a second string can be generated by globally replacing the # symbol with nothing to create the string ‘1 OR 2 OR 3 OR 4 …. OR 100’ to be used for those service providers where the search interface does not use the # symbol.

Figure 6.4.a: Combining concepts as search sets



6.4.8 Proximity operators (NEAR, NEXT and ADJ)

In some search interfaces it is necessary to specify, for example by using the 'NEXT' or 'ADJ' operator, that two search terms should be adjacent to each other, as the search might simply default to finding both words in the document as if the 'AND' operator had been used. It should be noted that the 'NEXT' operator in *The Cochrane Library* is more sensitive (i.e. retrieves more hits) than the alternative method of phrase searching using quotation marks, since quotation marks specify that exact phrase whereas the 'NEXT' operator incorporates auto-pluralization and auto-singularization as well as other variant word endings.

In addition, it is possible in many search interfaces to specify that the words should be within a specific number of words of each other. For example, the 'NEAR' operator in *The Cochrane Library* will find the search terms within six words of each other. This results in higher sensitivity than simple phrase searching or use of the 'NEXT' operator but greater precision than use of the 'AND' operator. It is, therefore, desirable to use this operator where available and relevant.

6.4.9 Language, date and document format restrictions

Research related to identifying trials has recently focused on the effect of excluding versus including from meta-analyses trials reported in languages other than English. This question is particularly important because the identification and translation of, or at least data extraction from, trials reported

in languages other than English can substantially add to the costs of a review and the time taken to complete it. For further discussion of these issues, see Chapter 10 (Section 10.2.2.4). Whenever possible review authors should attempt to identify and assess for eligibility all possibly relevant reports of trials irrespective of language of publication. No language restrictions should be included in the search strategy. Date restrictions should be applied only if it is known that relevant studies could only have been reported during a specific time period, for example if the intervention was only available after a certain time point. Format restrictions such as excluding letters are not recommended because letters may contain important additional information relating to an earlier trial report or new information about a trial not reported elsewhere.

6.4.10 Identifying fraudulent studies, other retracted publications, errata and comments

When considering the eligibility of studies for inclusion in a Cochrane review, it is important to be aware that some studies may have been found to be fraudulent or may for other reasons have been retracted since publication. Reports of studies indexed in MEDLINE that have been retracted (as fraudulent or for other reasons) will have the Publication Type term 'Retracted Publication' added to the record. The article giving notice of the retraction will have the Publication Type term 'Retraction of Publication' assigned. Prior to any decision being taken to retract an article, articles may be published that refer to an original article and raise concerns of this sort. Such articles would be classified as a Comment. The US National Library of Medicine's (NLM's) policy on this is that "Among the types of articles that will be considered comments are: ….. announcements or notices that report questionable science or investigations of scientific misconduct (sometimes published as 'Expression of concern')".

- www.nlm.nih.gov/pubs/factsheets/errata.html

In addition, articles may have been partially retracted, corrected through a published erratum or may have been corrected and re-published in full. When updating a review, it is important to search MEDLINE for the latest version of the citations to the records for the included studies. In some display formats of some versions of MEDLINE the retracted publication, erratum and comment statements are included in the citation data immediately after the title and are, therefore, highly visible. This is not, however, always the case so care should be taken to ensure that this information is always retrieved in all searches by downloading the appropriate fields together with the citation data (see Section 6.5.2). For further details of NLM's policy and practice in this area see:

- www.nlm.nih.gov/pubs/factsheets/errata.html

6.4.11 Search filters

Search filters are search strategies that are designed to retrieve specific types of records, such as those of a particular methodological design. They may be subjectively derived strategies such as the original Cochrane Highly Sensitive Search Strategy for identifying reports of randomized trials in MEDLINE (Dickersin 1994) or they may be objectively derived by word frequency analysis and tested on data sets of relevant records to assess their sensitivity and precision, such as the search strategies below for identifying randomized trials in MEDLINE (Glanville 2006). Recently a search filters web site has been developed by the UK InterTASC Information Specialists Subgroup (ISSG), which is the group of information professionals supporting research groups within England and Scotland providing technology assessments to the National Institute for Health and Clinical Excellence (NICE) (Glanville 2008). The purpose of the web site is to list methodological search filters and to provide critical appraisals of the various filters. The site includes, amongst others, filters for identifying systematic reviews, randomized and non-randomized studies and qualitative research in a range of databases and across a range of service providers.

- www.york.ac.uk/inst/crd/intertasc/

Search filters should be used with caution. They should be assessed not only for the reliability of their development and reported performance but also for their current accuracy, relevance and effectiveness given the frequent interface and indexing changes affecting databases.

The ISSG offer a search filter appraisal tool to assist with assessing search filters and examples can be seen on the website.

- www.york.ac.uk/inst/crd/intertasc/qualitat.htm

6.4.11.1 The Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE

The first Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE was designed by Carol Lefebvre and published in 1994 (Dickersin 1994). This strategy was subsequently published in the *Handbook* and has been adapted and updated as necessary over time. The Cochrane Highly Sensitive Search Strategies for MEDLINE in subsequent sections are adapted from strategies first published in 2006 as a result of a frequency analysis of MeSH terms and free-text terms occurring in the titles and abstracts of MEDLINE-indexed records of reports of randomized controlled trials (Glanville 2006), using methods of search strategy design first developed by the authors to identify systematic reviews in MEDLINE (White 2001).

Two strategies are offered: a sensitivity-maximizing version and a sensitivity- and precision-maximizing version. It is recommended that searches for trials for inclusion in Cochrane reviews begin with the sensitivity-maximizing version in combination with a highly sensitive subject search. If this retrieves an unmanageable number of references the sensitivity- and precision-maximizing version should be used instead. It should be borne in mind that MEDLINE abstracts can be read quite quickly as they are relatively short and, at a conservative estimate of 30 seconds per abstract, 1000 abstracts can be read in approximately 8 hours.

The strategies have been updated, after re-analysis of the data used to derive those strategies, to reflect changes in indexing policy introduced by the US National Library of Medicine since the original analysis and changes in search syntax. These changes include:

- no longer assigning ‘Clinical Trial’ as a Publication Type to all records indexed with ‘Randomized Controlled Trial’ or ‘Controlled Clinical Trial’ as a Publication Type; and
- the change of the MeSH term CLINICAL TRIALS to CLINICAL TRIALS AS TOPIC.

The strategies are given in Box 6.4.a and Box 6.4.b for PubMed, and in Box 6.4.c and Box 6.4.d for Ovid.

It must be borne in mind that the strategies below are based on data derived from MEDLINE-indexed records and were designed to be run in MEDLINE. These strategies are not designed to retrieve ‘in process’ and other records not indexed with MeSH. It is, therefore, recommended that these strategies are run in the MEDLINE-indexed versions of MEDLINE and separate searches for non-indexed records are run in the database containing the ‘in process’ and non-indexed records. For example, in Ovid the strategies below should be run and updated in databases such as ‘Ovid MEDLINE(R) 1950 to Month Week X 200X’ and non-indexed records should be searched for in ‘Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations Month X, 200X’. For identifying non-indexed records a range of truncated free-text terms would be required, such as random, placebo, trial, etc, and the search must not be limited to humans (as the records are not yet indexed as humans).

As discussed in Section 6.3.2.1, MEDLINE has been searched from 1966 to 2004 inclusive, using previous versions of the Cochrane Highly Sensitive Search Strategy for identifying randomized trials, and records of reports of trials (on the basis of the titles and abstracts only) have been re-indexed in

MEDLINE and included in CENTRAL. Refer to Section 6.3.2.1 and 6.3.3.2 for further guidance as to the appropriate use of these Highly Sensitive Search Strategies.

Box 6.4.a: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); PubMed format

#1	randomized controlled trial [pt]
#2	controlled clinical trial [pt]
#3	randomized [tiab]
#4	placebo [tiab]
#5	drug therapy [sh]
#6	randomly [tiab]
#7	trial [tiab]
#8	groups [tiab]
#9	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
#10	animals [mh] NOT humans [mh]
#11	#9 NOT #10

PubMed search syntax

[pt] denotes a Publication Type term;

[tiab] denotes a word in the title or abstract;

[sh] denotes a subheading;

[mh] denotes a Medical Subject Heading (MeSH) term ('exploded');

[mesh: noexp] denotes a Medical Subject Heading (MeSH) term (not 'exploded');

[ti] denotes a word in the title.

Box 6.4.b: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); PubMed format

#1	randomized controlled trial [pt]
#2	controlled clinical trial [pt]
#3	randomized [tiab]
#4	placebo [tiab]
#5	clinical trials as topic [mesh: noexp]
#6	randomly [tiab]
#7	trial [ti]
#8	#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7
#9	animals [mh] NOT humans [mh]
#10	#8 NOT #9

The search syntax is explained in Box 6.4.a.

Box 6.4.c: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); Ovid format

1	randomized controlled trial.pt.
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.
5	drug therapy.fs.
6	randomly.ab.
7	trial.ab.
8	groups.ab.
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10	exp animals/ not humans.sh.
11	9 not 10

Ovid search syntax

.pt. denotes a Publication Type term;

.ab. denotes a word in the abstract;

.fs. denotes a 'floating' subheading;

.sh. denotes a Medical Subject Heading (MeSH) term;

.ti. denotes a word in the title.

Box 6.4.d: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); Ovid format

1	randomized controlled trial.pt.
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.
5	clinical trials as topic.sh.
6	randomly.ab.
7	trial.ti.
8	1 or 2 or 3 or 4 or 5 or 6 or 7
9	exp animals/ not humans.sh.
10	8 not 9

The search syntax is explained in Box 6.4.c.

6.4.11.2 Search filters for identifying randomized trials in EMBASE

The UK Cochrane Centre is working on designing an objectively derived highly sensitive search strategy for identifying reports of randomized trials in EMBASE, using word frequency analysis methods similar to those used to design the highly sensitive search strategies for identifying randomized trials in MEDLINE described in Section 6.4.11.1 (Glanville 2006). Review authors wishing to conduct their own searches of EMBASE in the meanwhile might wish to consider using the search terms listed in Section 6.3.2.2 that are currently used by the UK Cochrane Centre to identify

EMBASE reports of randomized trials for inclusion in CENTRAL (Lefebvre 2008). Alternatively, the search filter designed by Wong and colleagues for identifying what they define as "clinically sound treatment studies" in EMBASE may be used (Wong 2006).

As discussed in Section 6.3.2.2, EMBASE has been searched from 1980 to 2006 inclusive, using the terms listed in that section, and records of reports of trials (on the basis of the titles and abstracts only) have been included in CENTRAL.

6.4.12 Updating searches

When a Cochrane review is updated, the search process (i.e. deciding which databases and other sources to search for which years) will have to be reviewed. Those databases that were previously searched and are considered relevant for the update will need to be searched again. The previous search strategies will need to be updated to reflect issues such as: changes in indexing such as the addition or removal of controlled vocabulary terms (MeSH, EMTREE etc); changes in search syntax; comments or criticisms of the previous search strategies. If any of the databases originally searched are not to be searched for the update this should be explained and justified. New databases or other sources may have been produced or become available to the review author or Trials Search Co-ordinator and these should also be considered.

Caution should be exercised with the use of update limits when searching across MEDLINE-indexed and un-indexed records simultaneously such as in PubMed or in the Ovid MEDLINE 'In-Process & Other Non-Indexed Citations and Ovid MEDLINE 1950 to Present' file. Where possible, separate files should be selected and searched separately, such as the Ovid MEDLINE '1950 to Month Week X 200X', and the non-indexed records should be searched for in the Ovid MEDLINE 'In-Process & Other Non-Indexed Citations Month X, 200X' file. For further guidance on this issue, contact a Trials Search Co-ordinator.

6.4.13 Demonstration search strategies

Box 6.4.e provides a demonstration search strategy for CENTRAL for the topic 'Tamoxifen for breast cancer'. Note that it includes topic terms only (a randomized trial filter is not appropriate for CENTRAL). There is no limiting to humans only. The strategy is provided for illustrative purposes only: searches of CENTRAL for studies to include in a systematic review would have many more search terms for each of the concepts.

Box 6.4.f provides a demonstration search strategy for MEDLINE (Ovid format) for the topic 'Tamoxifen for breast cancer'. Note that both topic terms and a randomized trial filter are used for MEDLINE. The search is limited to humans. The strategy is provided for illustrative purposes only: searches of MEDLINE for systematic reviews would have many more search terms for each of the concepts

Box 6.4.e: Demonstration search strategy for CENTRAL, for the topic 'Tamoxifen for breast cancer'

#1	MeSH descriptor Breast Neoplasms explode all trees
#2	breast near cancer*
#3	breast near neoplasm*
#4	breast near carcinoma*

#5	breast near tumour*
#6	breast near tumor*
#7	#1 OR #2 OR #3 OR #4 OR #5 OR #6
#8	MeSH descriptor Tamoxifen explode all trees
#9	tamoxifen
#10	#8 OR #9
#11	#7 AND #10

The 'near' operator defaults to within six words;

'*' indicates truncation.

Box 6.4.f: Demonstration search strategy for MEDLINE (Ovid format), for the topic 'Tamoxifen for breast cancer'

1	randomized controlled trial.pt.
2	controlled clinical trial.pt.
3	randomized.ab.
4	placebo.ab.
5	drug therapy.fs.
6	randomly.ab.
7	trial.ab.
8	groups.ab.
9	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10	animals.sh. not (humans.sh. and animals.sh.)
11.	9 not 10
12.	exp Breast Neoplasms/
13.	(breast adj6 cancer$).mp.
14.	(breast adj6 neoplasm$).mp.
15.	(breast adj6 carcinoma$).mp.
16.	(breast adj6 tumour$).mp.
17.	(breast adj6 tumor$).mp.
18.	12 or 13 or 14 or 15 or 16 or 17
19.	exp Tamoxifen/
20.	tamoxifen.mp.
21.	19 or 20
22.	11 and 18 and 21

The 'adj6' operator indicates within six words;

'$' indicates truncation;

.mp. indicates a search of title, original title, abstract, name of substance word and subject heading word.

6.4.14 Summary points

- Cochrane review authors should contact their Trials Search Co-ordinator *before* starting a search.
- For most Cochrane reviews, the search structure in most databases will be comprised of a subject search for population or condition and intervention together with a methodological filter for the study design, such as randomized trials.
- For searches of CENTRAL, do not apply a randomized trial filter and do not limit to human.
- Avoid too many *different* search concepts but use a wide variety of synonyms and related terms (both free text and controlled vocabulary terms) combined with 'OR' within *each* concept.
- Combine different concepts with 'AND'.
- Avoid use of the 'NOT' operator in combining search sets.
- Aim for high sensitivity and be prepared to accept low precision.
- Do not apply language restrictions to the search strategy.
- Searches designed for a specific database and service provider will need to be 'translated' for use in another database or service provider.
- Ensure awareness of any retracted publications (e.g. fraudulent publications), errata and comments.
- For identifying randomized trials in MEDLINE, begin with the sensitivity-maximizing version of the Cochrane Highly Sensitive Search Strategy. If this retrieves an unmanageable number of references, use the sensitivity- and precision-maximizing version instead.
- For update searches, where possible, separate database files should be selected and searched separately for the MEDLINE-indexed records and the non-indexed in-process records.

6.5 Managing references

6.5.1 Bibliographic software

Specially designed bibliographic or reference management software such as EndNote, ProCite, Reference Manager and RefWorks is useful and relatively easy to use to keep track of references to and reports of studies. The choice of which software to use is likely to be influenced by what is available and thus supported at the review author's institution. For a comparison of the above products and links to reviews of other bibliographic software packages see:

- www.burioni.it/forum/dellorso/bms-dasp/text/

Of the packages listed above, ProCite is generally considered to be very efficient for identifying duplicate references but is no longer updated by the suppliers. It does not support the wider range of character sets allowing references to be entered correctly in languages other than English, whereas EndNote does. Bibliographic software also facilitates storage of information about the methods and process of a search. For example, separate unused fields can be used to store information such as 1) the name of the database or other source details from which a trial report was identified, 2) when and from where an article was ordered and the date of article receipt and 3) whether the study associated with an article was included in or excluded from a review and, if excluded, the reasons for exclusion.

Files for importing references from CENTRAL into bibliographic software are available from the Cochrane Information Retrieval Methods Group web site at:

- www.cochrane.org/docs/import.htm

6.5.2 Which fields to download

In addition to the full record citation a number of key fields should be considered for downloading from databases where they are available. Further detailed guidance on which fields to download has been compiled by the Trials Search Co-ordinators' Working Group and is available in a document entitled 'TSC User Guide to Managing Specialized Registers and Handsearch Records' at:

- www.cochrane.org/resources/hsearch.htm

Abstract: abstracts can be used to eliminate clearly irrelevant reports, obviating the need to obtain the full text of those reports or to return to the bibliographic database at a later time.

Accession number / unique identifier: it is advisable to set aside an unused field for storing the unique identifier / accession number of records downloaded, such as the PubMed ID number (PMID). This allows subsequent linkage to the full database record and also facilitates information management such as duplicate detection and removal.

Affiliation / address: may include the institutional affiliation and / or e-mail address of the author(s).

Article identifier / digital object identifier (DOI): can be used to cite and link to the full record.

Clinical trial number: if the record contains a clinical trial number such as those assigned by the ClinicalTrials.gov or ISRCTN schemes or a number allocated by the sponsor of the trial, these should be downloaded to aid linking of trial reports to the original studies. An example of this is the Clinical Trial Number (CN) field recently introduced in EMBASE.

Index terms / thesaurus terms / keywords: see Section 6.4.5. These help indicate why records were retrieved if the title and abstract lack detail.

Language: language of publication of the original article.

Comments, corrections, errata, retractions and updates: it is important to ensure that any fields that relate to subsequently published comments, corrections, errata, retractions and updates are selected for inclusion in the download, so that any impact of these subsequent publications can be taken into account. The most important fields to consider, together with their field labels in PubMed, are provided in Box 6.5.a.

- www.nlm.nih.gov/bsd/mms/medlineelements.html#cc

Box 6.5.a: Important field labels in PubMed

CIN: 'Comment in'

CON: 'Comment on'
CRI: 'Corrected and republished in'

CRF: ''Corrected and republished from'

EIN: 'Erratum in'
EFR: 'Erratum for'

PRIN: 'Partial retraction in'
PROF: 'Partial retraction of'

RIN: 'Retraction in'
ROF: 'Retraction of'

RPI: 'Republished in'

RPF: ''Republished from'

UIN: 'Update in'
UOF: 'Update of'

6.5.3 Summary points

- Use bibliographic software to manage references.

- Ensure that all the necessary fields are downloaded.

6.6 Documenting and reporting the search process

6.6.1 Documenting the search process

The search process needs to be documented in enough detail throughout the process to ensure that it can be reported correctly in the review, to the extent that all the searches of all the databases are reproducible. It should be borne in mind at the outset that the full search strategies for each database will need to be included in an Appendix of the review. The search strategies will need to be copied and pasted exactly as run and included in full, together with the search set numbers and the number of records retrieved. The number of records retrieved will need to be recorded in the Results section of the review, under the heading 'Results of the search' (see Chapter 4, Section 4.5). The search strategies should not be re-typed as this can introduce errors. A recent study has shown lack of compliance with guidance in the *Handbook* with respect to search strategy description in Cochrane reviews (Sampson 2006). In the majority of CRGs, the Trials Search Co-ordinators are now asked to comment on the search strategy sections of a review as part of the sign-off process prior to a review being considered ready for publication in the *CDSR*. It is, therefore, recommended that review authors should seek guidance from their Trials Search Co-ordinator at the earliest opportunity with respect to documenting the process to facilitate writing up this section of the review. As mentioned elsewhere in this chapter, it is particularly important to save locally or file print copies of any information found on the internet, such as information about ongoing trials, as this information may no longer be accessible at the time the review is written up.

6.6.2 Reporting the search process

6.6.2.1 Reporting the search process in the protocol

The inclusion of any search strategies in the protocol for a Cochrane review is optional. Where searches have already been undertaken at the protocol stage it is considered useful to include them in the protocol so that they can be commented upon in the same way as other aspects of the protocol. Some CRGs are of the view that no searches should be undertaken until the protocol is finalized for publication as knowledge of the available studies might influence aspects of the protocol such as inclusion criteria.

6.6.2.2 Reporting the search process in the review

Reporting the search process in the review abstract

- List all databases searched.
- Note the dates of the last search for each database or the period searched.
- Note any language or publication status restrictions (but refer to Section 6.4.9).
- List individuals or organizations contacted.

For further guidance on how this information should be listed see Chapter 11 (Section 11.8).

Reporting the search process in the Methods section

In the 'Search methods for identification of studies' section(s):

- List all databases searched.
- Note the dates of the last search for each database AND the period searched.
- Note any language or publication status restrictions (but refer to Section 6.4.9).

- List grey literature sources.
- List individuals or organizations contacted.
- List any journals and conference proceedings specifically handsearched for the review.
- List any other sources searched (e.g. reference lists, the internet).

The full search strategies for each database should be included in an Appendix of the review to avoid interrupting the flow of the text of the review. The search strategies should be copied and pasted exactly as run and included in full together with the line numbers for each search set. They should not be re-typed as this can introduce errors. For further detailed guidance on this contact the Trials Search Co-ordinator.

Reporting the search process in the Results section

The number of hits retrieved by the electronic searches should be included in the Results section.

Reporting date of the search

A single date should be specified in the 'Date of search' field, to indicate when the most recent comprehensive search was started. For more information on specifying this date, see Chapter 3 (Section 3.3.3).

6.6.3 Summary points

- Seek guidance on documenting the search process from a Trials Search Co-ordinator before starting searching.
- The full strategy for each search of each database should be copied and pasted into an Appendix of the review.
- The total number of hits retrieved by each search strategy should be included in the Results section.
- Save locally or file print copies of any information found on the internet, such as information about ongoing trials.
- Refer to Chapter 4 (Section 4.5) and Chapter 11 (Section 11.8) for more information on what to report in the review and the abstract, respectively.

6.7 Chapter information

Authors: Carol Lefebvre, Eric Manheimer and Julie Glanville on behalf of the Cochrane Information Retrieval Methods Group.

This chapter should be cited as: Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. Available from www.cochrane-handbook.org.

Acknowledgements: This chapter has been developed from sections of previous editions of the *Handbook* co-authored since 1995 by Kay Dickersin, Kristen Larson, Carol Lefebvre and Eric Manheimer. Many of the sources listed in this chapter have been brought to our attention by a variety of people over the years and we should like to acknowledge this. We should like to thank the information specialists who shared with us information and documentation about their search processes. We should also like to thank Cochrane Trials Search Co-ordinators, members of the Cochrane Information Retrieval Methods Group (see Box 6.7.a), the Health Technology Assessment

International Special Interest Group on Information Resources and the InterTASC Information Specialists' Subgroup for comments on earlier drafts of this Chapter, Anne Eisinga for proof-reading the search strategies, and the two peer reviewers, Steve McDonald and Ruth Mitchell, for their detailed and constructive comments.

Box 6.7.a: The Cochrane Information Retrieval Methods Group

The Information Retrieval Methods Group (IRMG) aims to provide advice and support, to conduct research and to facilitate information exchange regarding methods to support the information retrieval activities of The Cochrane Collaboration. The group was officially registered with the Collaboration in November 2004. Members concentrate on providing practical support for the development of information retrieval techniques and facilities for information searchers. The group's aims are realized by the following activities:

- Offering advice on information retrieval policy and practice;
- Providing training and support;
- Conducting empirical research (including systematic reviews) into information retrieval methods;
- Helping to monitor the quality of searching techniques employed in systematic reviews;
- Liaising with members of the Campbell Collaboration to avoid duplication of effort in areas of information retrieval of interest to both the Cochrane and Campbell Collaborations;
- Serving as a forum for discussion.

Web site: www.cochrane.org/docs/irmg.htm

6.8 References

Bennett 2003
Bennett DA, Jull A. FDA: untapped source of unpublished trials. *The Lancet* 2003; 361: 1402-1403.

De Angelis 2004
De Angelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, Kotzin S, Laine C, Marusic A, Overbeke AJ, Schroeder TV, Sox HC, Van der Weyden MB, International Committee of Medical Journal Editors. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA* 2004; 292: 1363-1364.

De Angelis 2005
De Angelis CD, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, Kotzin S, Laine C, Marusic A, Overbeke AJ, Schroeder TV, Sox HC, Van der Weyden MB, International Committee of Medical Journal Editors. Is this clinical trial fully registered? A statement from the International Committee of Medical Journal Editors. *JAMA* 2004; 293: 2927-2929.

Dickersin 1994
Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994; 309: 1286-1291.

Dickersin 2002
Dickersin K, Manheimer E, Wieland S, Robinson KA, Lefebvre C, McDonald S, CENTRAL Development Group. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Evaluation and the Health Professions* 2002; 25: 38-64.

Eisinga 2007
Eisinga A, Siegfried N, Clarke M. The sensitivity and precision of search terms in Phases I, II and III of the Cochrane Highly Sensitive Search Strategy for identifying reports of randomized trials in

MEDLINE in a specific area of health care - HIV/AIDS prevention and treatment interventions. *Health Information and Libraries Journal* 2007; 24: 103-109.

Eysenbach 2001
Eysenbach G, Tuische J, Diepgen TL. Evaluation of the usefulness of Internet searches to identify unpublished clinical trials for systematic reviews. *Medical Informatics and the Internet in Medicine* 2001; 26: 203-218.

Glanville 2006
Glanville JM, Lefebvre C, Miles JN, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: ten years on. *Journal of the Medical Library Association* 2006; 94: 130-136.

Glanville 2008
Glanville J, Bayliss S, Booth A, Dundar Y, Fleeman ND, Foster L, Fraser C, Fernandes H, Fry-Smith A, Golder S, Lefebvre C, Miller C, Paisley S, Payne L, Price AM, Welch K, InterTASC Information Specialists' Subgroup. So many filters, so little time: The development of a Search Filter Appraisal Checklist. *Journal of the Medical Library Association* (in press, 2008).

Golder 2006
Golder S, McIntosh HM, Duffy S, Glanville J, Centre for Reviews and Dissemination and UK Cochrane Centre Search Filters Design Group. Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. *Health Information and Libraries Journal* 2006; 23: 3-12.

Greenhalgh 2005
Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ* 2005; 331: 1064-1065.

Hetherington 1989
Hetherington J, Dickersin K, Chalmers I, Meinert CL. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989; 84: 374-380.

Hopewell 2007a
Hopewell S, Clarke M, Lefebvre C, Scherer R. Handsearching versus electronic searching to identify reports of randomized trials. *Cochrane Database of Systematic Reviews* 2007, Issue 2. Art No: MR000001.

Hopewell 2007b
Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database of Systematic Reviews* 2007, Issue 2. Art No: MR000010.

Horton 1997
Horton R. Medical editors trial amnesty. *The Lancet* 1997; 350: 756.

Khan 2001
Khan KS, ter Riet G, Glanville J, Sowden AJ, Kleijnen J (editors). *Undertaking Systematic Reviews of Research on Effectiveness: CRD's Guidance for those Carrying Out or Commissioning Reviews (CRD Report Number 4)* (2nd edition). York (UK): NHS Centre for Reviews and Dissemination, University of York, 2001.

Lefebvre 2001
Lefebvre C, Clarke M. Identifying randomised trials. In: Egger M, Davey Smith G, Altman DG (editors). *Systematic Reviews in Health Care: Meta-analysis in Context* (2nd edition). London (UK): BMJ Publication Group, 2001.

Lefebvre 2008
Lefebvre C, Eisinga A, McDonald S, Paul N. Enhancing access to reports of clinical trials published world-wide - the contribution of EMBASE records to the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library. *Emerging Themes in Epidemiology* (in press, 2008).

MacLean 2003
MacLean CH, Morton SC, Ofman JJ, Roth EA, Shekelle PG. How useful are unpublished data from the Food and Drug Administration in meta-analysis? *Journal of Clinical Epidemiology* 2003; 56: 44-51.

Mallett 2002
Mallett S, Hopewell S, Clarke M. Grey literature in systematic reviews: The first 1000 Cochrane systematic reviews. *Fourth Symposium on Systematic Reviews: Pushing the Boundaries*, Oxford (UK), 2002.

Manheimer 2002
Manheimer E, Anderson D. Survey of public information about ongoing clinical trials funded by industry: evaluation of completeness and accessibility. *BMJ* 2002; 325: 528-531.

McDonald 2002
McDonald S. Improving access to the international coverage of reports of controlled trials in electronic databases: a search of the Australasian Medical Index. *Health Information and Libraries Journal* 2002; 19: 14-20.

Montori 2005
Montori VM, Wilczynski NL, Morgan D, Haynes RB. Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ* 2005; 330: 68.

Royle 2003
Royle P, Milne R. Literature searching for randomized controlled trials used in Cochrane reviews: rapid versus exhaustive searches. *International Journal of Technology Assessment in Health Care* 2003; 19: 591-603.

Sampson 2006
Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. *Journal of Clinical Epidemiology* 2006; 59: 1057-1063.

Scherer 2007
Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2007, Issue 2. Art No: MR000005.

Suarez-Almazor 2000
Suarez-Almazor ME, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Controlled Clinical Trials* 2000; 21: 476-487.

White 2001
White VJ, Glanville JM, Lefebvre C, Sheldon TA. A statistical approach to designing search filters to find systematic reviews: objectivity enhances accuracy. *Journal of Information Science* 2001; 27: 357-370.

Whiting 2008
Whiting P, Westwood M, Burke M, Sterne J, Glanville J. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *Journal of Clinical Epidemiology* 2008; 61: 357.e1-357.e10.

Wilczynski 2007
Wilczynski NL, Haynes RB, Hedges Team. EMBASE search strategies achieved high sensitivity and specificity for retrieving methodologically sound systematic reviews. *Journal of Clinical Epidemiology* 2007; 60: 29-33.

Wong 2006
Wong SS, Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound treatment studies in EMBASE. *Journal of the Medical Library Association* 2006; 94: 41-47.

This is an archived version of the Handbook.
For the current version, please go to cochrane.org/handbook

Chapter 13: Including non-randomized studies

Authors: Barnaby C Reeves, Jonathan J Deeks, Julian PT Higgins and George A Wells on behalf of the Cochrane Non-Randomised Studies Methods Group.

This extract is from *Handbook* version 5.0.2. For guidance on how to cite it, see Section 13.8. The material is also published in Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions* (ISBN 978-0470057964) by John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England, Telephone (+44) 1243 779777; Email (for orders and customer service enquiries): cs-books@wiley.co.uk. Visit their Home Page on www.wiley.com.

Key points

- For some Cochrane reviews, the question of interest cannot be answered by randomized trials, and review authors may be justified in including non-randomized studies.
- Potential biases are likely to be greater for non-randomized studies compared with randomized trials, so results should always be interpreted with caution when they are included in reviews and meta-analyses. Particular concerns arise with respect to differences between people in different intervention groups (selection bias) and studies that do not explicitly report having had a protocol (reporting bias).
- We recommend that eligibility criteria, data collection and critical assessment of included studies place an emphasis on specific features of study design (e.g. which parts of the study were prospectively designed) rather than 'labels' for study designs (such as case-control versus cohort).
- Risk of bias in non-randomized studies can be assessed in a similar manner to that used for randomized trials, although more attention must be paid to the possibility of selection bias.
- Meta-analyses of non-randomized studies must consider how potential confounders are addressed, and consider the likelihood of increased heterogeneity resulting from residual confounding and from other biases that vary across studies.

13.1 Introduction

13.1.1 What this chapter is about

This chapter has been prepared by the Non-Randomised Studies Methods Group (NRSMG) of The Cochrane Collaboration (see Box 13.8.a). It is intended to support review authors who are considering including non-randomized studies in Cochrane reviews. **Non-randomized studies** (NRS) are defined

here as any quantitative study estimating the effectiveness of an intervention (harm or benefit) that does not use randomization to allocate units to comparison groups. This includes studies where allocation occurs in the course of usual treatment decisions or peoples' choices, i.e. studies usually called *observational*. There are many types of non-randomized intervention study, including cohort studies, case-control studies, controlled before-and-after studies, interrupted-time-series studies and controlled trials that use inappropriate randomization strategies (sometimes called quasi-randomized studies). Box 13.1.a summarizes some commonly-used study design labels for non-randomized studies. We explain in Section 13.5.1 why we do not necessarily advise that these labels are used in Cochrane reviews.

This chapter aims to describe the particular challenges that arise if NRS are included in a Cochrane review, and is informed by theoretical or epidemiological considerations, empirical research, and discussions among members of the NRSMG. The chapter makes recommendations about what to do when it is possible to support the recommendations on the basis of evidence or established theory. When it is not possible to make any recommendations, the chapter aims to set out the pros and cons of alternative actions and to identify questions for further methodological research.

Review authors who are considering including NRS in a Cochrane review should not start with this chapter unless they are already familiar with the process of preparing a systematic review of randomized trials. The format and basic steps of a Cochrane review should be the same whether it includes only randomized trials or includes NRS. The reader is referred to Part 1 of the Handbook for a detailed description of these steps. Every step in carrying out a systematic review is more difficult when NRS are included and a review author should seek to include expert epidemiologists and methodologists in the review team. As an example of such collaboration, a review of NRS included nine authors, five of whom were methodologists (Siegfried 2003).

Box 13.1.a: Some types of NRS design used for evaluating the effects of interventions

Designs are distinguished below by labels in common use and descriptions are intentionally non-specific because the labels are interpreted in different ways with respect to details. The NRSMG does not advocate using these labels for reasons explained in Section 13.5.1.

Non-randomized controlled trial.	An experimental study in which people are allocated to different interventions using methods that are not random.
Controlled before-and-after study.	A study in which observations are made before and after the implementation of an intervention, both in a group that receives the intervention and in a control group that does not.
Interrupted-time-series study.	A study that uses observations at multiple time points before and after an intervention (the 'interruption'). The design attempts to detect whether the intervention has had an effect significantly greater than any underlying trend over time.
Historically controlled study.	A study that compares a group of participants receiving an intervention with a similar group from the past who did not.
Cohort study.	A study in which a defined group of people (the cohort) is followed over time, to examine associations between different interventions received and subsequent outcomes. A 'prospective' cohort study recruits participants before any intervention and follows them into the future. A 'retrospective' cohort study identifies subjects from past records describing the interventions received and follows them from the time of those records.

Case-control study.	A study that compares people with a specific outcome of interest ('cases') with people from the same source population but without that outcome ('controls'), to examine the association between the outcome and prior exposure (e.g. having an intervention). This design is particularly useful when the outcome is rare.
Cross-sectional study.	A study that collects information on interventions (past or present) and current health outcomes, i.e. restricted to health states, for a group of people at a particular point in time, to examine associations between the outcomes and exposure to interventions.
Case series (uncontrolled longitudinal study).	Observations are made on a series of individuals, usually all receiving the same intervention, before and after an intervention but with no control group.

13.1.2 Why consider non-randomized studies?

The Cochrane Collaboration focuses particularly on systematic reviews of randomized trials because they are more likely to provide unbiased information than other study designs about the differential effects of alternative forms of health care. Reviews of NRS are only likely to be undertaken when the question of interest cannot be answered by a review of randomized trials. The NRSMG believes that review authors may be justified in including NRS which are moderately susceptible to bias. Broadly, the NRSMG considers that there are three main reasons for including NRS in a Cochrane review:

a) To examine the case for undertaking a randomized trial by providing an explicit evaluation of the weaknesses of available NRS. The findings of a review of NRS may also be useful to inform the design of a subsequent randomized trial, e.g. through the identification of relevant subgroups.

b) To provide evidence of the effects (benefit or harm) of interventions that cannot be randomized, or which are extremely unlikely to be studied in randomized trials. In these contexts, a disinterested (free from bias and partiality) review that systematically reports the findings and limitations of available NRS can be useful.

c) To provide evidence of effects (benefit or harm) that cannot be adequately studied in randomized trials, such as long-term and rare outcomes, or outcomes that were not known to be important when existing, major randomized trials were conducted.

Three other reasons are often cited in support of systematic reviews of NRS but are poor justifications:

d) Studying effects in patient groups not recruited to randomized trials (such as children, pregnant women, the elderly). Although it is important to consider whether the results of trials can be generalized to people who are excluded from them, it is not clear that this can be achieved by consideration of non-randomized studies. Regardless of whether estimates from NRS agree or disagree with those of randomized trials, there is always potential for bias in the results of the NRS, such that misleading conclusions are drawn.

e) To supplement existing randomized trial evidence. Adding non-randomized to randomized evidence may change an imprecise but unbiased estimate into a precise but biased estimate, i.e. an exchange of undesirable uncertainty for unacceptable error.

f) When an intervention effect is really large. Implicitly, this is a result-driven or *post hoc* justification, since the review (or some other synthesis of the evidence) needs to be undertaken to observe the likely size of the effects. Whilst it is easier to argue that large effects are less likely to be completely explained by bias than small effects (Glasziou 2007), for the practice of health care it is still important to obtain unbiased estimates of the magnitude of large effects to make clinical and economic decisions (Reeves 2006). Thus randomized trials are still needed for large effects (and they need not be large if the effects are truly large). There may be ethical opposition to

randomized trials of interventions already suspected to be associated with a large benefit as a result of a systematic review of NRS, making it difficult to randomize participants, and interventions postulated to have large effects may also be difficult to randomize for other reasons (e.g. surgery vs. no surgery). However, the justification for a systematic review of NRS in these circumstances should be classified as (b), i.e. interventions that are unlikely to be randomized, rather than as (f).

13.1.3 Key issues about the inclusion of non-randomized studies in a Cochrane review

Randomized trials are the preferred design for studying the effects of healthcare interventions because, in most circumstances, the randomized trial is the study design that is least likely to be biased. Any Cochrane review must consider the risk of bias in individual primary studies, including both the likely direction and magnitude of bias (see Chapter 8). A review that includes NRS also requires review authors to do this. The principle of considering risk of bias is exactly the same. However, potential biases are likely to be greater for NRS compared with randomized trials. Review authors need to consider (a) the weaknesses of the designs that have been used (such as noting their potential to ascertain causality), (b) the execution of the studies through a careful assessment of their risk of bias, especially (c) the potential for selection bias and confounding to which all NRS are suspect and (d) the potential for reporting biases, including selective reporting of outcomes.

Susceptibility to selection bias (understood in this *Handbook* to mean differences in the baseline characteristics of individuals in different intervention groups, rather than whether the selected sample is representative of the population) is widely regarded as the principal difference between randomized trials and NRS. Randomization with adequate allocation sequence concealment reduces the possibility of systematic selection bias in randomized trials so that differences in characteristics between groups can be attributed to chance. In NRS, allocation to groups depends on other factors, often unknown. Confounding occurs when selection bias gives rise to imbalances between intervention and control groups (or case and control groups in case-control studies) on prognostic factors, i.e. the distributions of the factors differ between groups *and* the factors are associated with outcome. Confounding can have two effects in a meta-analysis: (a) shifting the estimate of the intervention effect (systematic bias) and (b) increasing the variability of the observed effects, introducing excessive heterogeneity among studies (Deeks 2003). It is important to consider both of these possible effects (see Section 13.6.1). Section 13.5 provides a more detailed discussion of susceptibility to bias in NRS.

13.1.4 The importance of a protocol for a Cochrane review that includes non-randomized studies

Chapter 2 establishes the importance of writing a protocol for a Cochrane review before carrying out the review. As the methodological choices made during a review of NRS are complex and may affect the review findings, a protocol is even more important for a review that includes NRS. The rationale for doing a review that includes NRS (see Section 13.1.2) should be documented in the protocol. The protocol should include much more detail than for a review of randomized trials, pre-specifying key methodological decisions about the methods to be used and the analyses that are planned. The protocol needs to specify details that are not relevant for randomized trials (e.g. the methods planned to identify potential confounding factors and to assess the susceptibility of primary studies to confounding), as well as providing more detail about standard steps in the review process that are more difficult when including NRS (e.g. specification of eligibility criteria and the search strategy for identifying eligible studies).

The NRSMG recognizes that it may not be possible to pre-specify all decisions about the methods used in a review. Nevertheless, review authors should aim to make all decisions about the methods for

the review without reference to the findings of primary studies, and report methodological decisions that had to be made or modified after collecting data about the study findings.

13.1.5 Structure of subsequent sections in the chapter

Each of the sections in this chapter, which focus in turn on different steps of the review process, is structured in the same way. First, for a particular step, we summarize what is different when NRS (compared with randomized trials) are included in Cochrane reviews and, where applicable, describe conceptual issues that need to be considered. This first part includes relevant evidence, where there is some. Second, we summarize our guidance and, where available, describe existing resources that are available to support review authors.

13.2 Developing criteria for including non-randomized studies

13.2.1 What is different when including non-randomized studies?

13.2.1.1 Including both randomized and non-randomized studies

Review authors may want to include NRS in a review because only a small number of randomized trials can be identified, or because of perceived limitations of the randomized trials. In this chapter, we strongly recommend that review authors should not make any attempt to combine evidence from randomized trials and NRS. This recommendation means that criteria for included study designs should generally specify randomized or non-randomized studies when trying to evaluate the effect of an intervention on a particular outcome. (However, a single review might consist of 'component' reviews that include different study designs for different outcomes, for example, randomized trials for evaluating benefits and NRS to evaluate harms; see Chapter 14.) Alternatively, where randomized trial evidence is desired but unlikely to be available, eligibility criteria could reasonably be structured to say that NRS would only be included where randomized trials are found not to be available. In time, as such a review is updated, the NRS may be dropped when randomized trials become available. Where both randomized trials and NRS of an intervention exist and, for one or more of the reasons given in Section 13.1.2, both are included in the review, these should be presented separately; alternatively, if there is an adequate number of randomized trials, comments about relevant NRS can be included in the Discussion section of a review although this is rarely particularly helpful.

13.2.1.2 Evaluating benefits and harms

Cochrane reviews aim to quantify the effects of healthcare interventions, both beneficial and harmful, and both expected and unexpected. Most reviews estimate the expected benefits of an intervention that are assessed in randomized trials. Randomized trials may report some of the harms of an intervention, either those which were expected and which the trial was designed to assess, or those which were not expected but which were collected in the trial as part of standard monitoring of safety. However, many serious harms of an intervention are too rare or do not appear during the follow-up period of randomized trials, and therefore will not be reported. Therefore, one of the most important roles for reviews of NRS is to assess potential unexpected or rare harms of interventions (reason (c) in Section 13.1.1). Criteria for selecting important and relevant studies for evaluating rare or long-term adverse and unexpected effects are difficult to set. Although the relative strengths and weaknesses of different study designs are the same as for beneficial outcomes, the choice of study designs to include may depend on both the frequency of an outcome and its importance. For example, for some rare adverse outcomes only case series or case-control studies may be available. Study designs that are more susceptible to bias may be acceptable for evaluation of serious events in the absence of better evidence.

Confounding may be less of a threat to the validity of a review when researching rare harms or unexpected effects of interventions than when researching expected effects, since it is argued that 'confounding by indication' mainly influences treatment decisions with respect to outcomes about which the clinicians are primarily concerned. However, confounding can never be ruled out because the same features that are confounders for the expected effects may also be direct confounders for the unexpected effects, or be correlated with features that are confounders.

A related issue is the need to distinguish between *quantifying* and *detecting* an effect of an intervention. Quantifying the intended benefits of an intervention – maximizing the precision of the estimate and minimizing susceptibility to bias – is critical when weighing up the relative merits of alternative interventions for the same condition. A review should also try to quantify the harms of an intervention, minimizing susceptibility to bias as far as possible. However, if a review can establish beyond reasonable doubt that an intervention causes a particular harm, the precision and susceptibility to bias of the estimated effect may not be critical. In other words, the seriousness of the harm may outweigh any benefit from the intervention. This situation is more likely to occur when there are competing interventions for a condition.

13.2.1.3 Determining which types of non-randomized study to include

A randomized trial is a prospective, experimental study design specifically involving random allocation of participants to interventions. Although there are variations in randomized trial design (including random allocation of individuals, clusters or body parts; multi-arm trials, factorial trials and cross-over trials) they constitute a distinctive study category. By contrast, NRS cover a number of fundamentally different designs, several of which were originally conceived in the context of aetiological epidemiology. Some of these are summarized in Box 13.1.a, although this is not an exhaustive list, and many studies combine ideas from different basic designs. As we discuss in 13.2.2 these labels are not consistently applied. The diversity of NRS designs raises two related questions. First, should all NRS designs of a particular effectiveness question be included in a review? Second, if review authors do not include all NRS designs, what criteria should be used to decide which study designs to include and which to exclude?

It is generally accepted that criteria should be set to limit the kinds of evidence included in a systematic review. The primary reason is that the risk of bias varies across studies. For this reason, many Cochrane reviews only include randomized trials (when available). For the same reason, it is argued that review authors should only include NRS that are least likely to be biased. It is not helpful to include primary studies in a review when the results of the studies are likely to be biased, even if there is no better evidence. This is because a misleading effect estimate may be more harmful to future patients than no estimate at all, particularly if the people using the evidence to make decisions are unaware of its limitations (Doll 1993, Peto 1995).

There is no agreement about the study design criteria that should be used to limit the inclusion of NRS in a Cochrane review. One strategy is to include only those study designs that will give reasonably valid effect estimates. Another strategy is to include the best available study designs which have been used to answer a question. The first strategy would mean that reviews are consistent and include the same types of NRS, but that some reviews include no studies at all. The second strategy leads to different reviews including different study designs according to what was available. For example, it might be entirely appropriate to use different criteria for inclusion when reviewing the harms, compared with the benefits, of an intervention. This approach is already evident in the *Cochrane Database of Systematic Reviews* (*CDSR*), with editors of some Cochrane Review Groups (CRGs) restricting reviews to randomized trials only and other CRG editors allowing specific types of NRS to be included in reviews (typically in healthcare areas where randomized trials are infrequent).

Whichever point of view is adopted, criteria can only be chosen with respect to a hierarchy of primary study designs, ranked in order of risk of bias according to study design features. Existing 'evidence hierarchies' for studies of effectiveness (Eccles 1996, National Health and Medical Research Council 1999, Oxford Centre for Evidence-based Medicine 2001) appear to have arisen largely by applying hierarchies for aetiological research questions to effectiveness questions. For example, cohort studies are conventionally regarded as providing better evidence than case-control studies. It is not clear that this is always appropriate since aetiological hierarchies place more emphasis on establishing causality (e.g. dose-response relationship, exposure preceding outcome) than on valid quantification of the effect size. Also, study designs used for studying the effects of interventions can be very much more diverse and complex (Shadish 2002) and may not be easily assimilated into existing evidence hierarchies (see the array of designs in Box 13.1.a, for example). Different designs are susceptible to different biases, and it is often unclear which biases have the greatest impact and how they vary between clinical situations.

13.2.1.4 Distinguishing between aetiology and effectiveness research questions

Including NRS in a Cochrane review allows, in principle, the inclusion of truly observational studies where the use of an intervention has occurred in the course of usual health care or daily life. For interventions that are not restricted to a medical setting, this may mean interventions that a study participant chooses to take, e.g. over-the-counter preparations. Including observational studies in a review also allows exposures to be studied that are not obviously 'interventions', e.g. nutritional choices, and other behaviours that may affect health. This introduces a 'grey area' between evidence about effectiveness and aetiology. It is important to distinguish carefully between different aetiological and effectiveness research questions related to a particular exposure. For example, nutritionists may be interested in the health-related effects of a diet that includes a minimum of five portions of fruit or vegetables per day ('five-a-day'), an aetiological question. On the other hand, public health professionals may be interested in the health-related effects of interventions to promote a change in diet to include 'five-a-day', an effectiveness question. Because of other differences between studies relevant to these two kinds of question (e.g. duration of follow-up and outcomes investigated), studies addressing the former type of question are often perceived as being 'better' or 'more relevant' without acknowledging or realizing that they are addressing different research questions. In other instances the health intervention being evaluated in the NRS will have been undertaken for a purpose other than improving health. For example, a review of circumcision for preventing transmission of HIV included NRS where circumcision had been undertaken for cultural or religious reasons (Siegfried 2003), and it was unclear whether using the intervention for health purposes would have the same effect.

13.2.2 Guidance and resources available to support review authors

Review authors should first check with the editors of the CRG under which they propose to register their protocol whether there is a CRG-specific policy in place about the inclusion of NRS in a review. Authors should also discuss with the editors the extent of methodological advice available in the CRG since they are likely to require more support than with a review that includes randomized trials only, and attempt to recruit informed methodologists to their review team. Regrettably, the NRSMG is not currently in a position to collaborate with authors on particular reviews, but encourages authors who include NRS in their reviews to feedback their experiences to the NRSMG, particularly where their experiences support, or contradict, the experiences described in this chapter.

Review authors intending to review the adverse effects (harms) of an intervention should read Chapter 14, which has been prepared by the Adverse Effects Methods Group.

We recommend that review authors use explicit study design features (NB: not study design labels) when deciding which types of NRS to include in a review. Members of the NRSMG have developed two lists that can be used for this purpose, although experience using them is limited. Table 13.2.a and

Table 13.2.b describe separate lists for individually-allocated and cluster-allocated studies. Sixteen (or fifteen) items are grouped under four headings:

1. Was there a comparison?
2. How were groups created?
3. Which parts of the study were prospective?
4. On which variables was comparability [between groups receiving different interventions] assessed?

The items are designed to characterize key features of studies which, on the basis of the experiences of NRSMG members and 'first principles' (rather than evidence), are suspected to define the major study design categories or to be associated with susceptibility to bias. The tables indicate which features are associated with different NRS designs, identified by labels that are more specific than those in Box 13.1.a. There is not total consensus about the use of these (column) labels. This disagreement does not mean that the row items are inappropriate or poorly described; the value of the lists depends on the agreement between review authors when classifying primary studies. We will also propose that these lists be used as checklists in the processes of data collection and as part of the critical assessment of the studies (Section 13.4.2 and Section 13.5.2). Instructions for using the items as checklists in Box 13.4.a provide further explanation of the terms.

A number of organizations are carrying out systematic reviews of NRS where there are no, or very few, randomized trials. Reviews are often commissioned on behalf of organizations responsible for issuing policy or guidance to healthcare professionals, e.g. the National Institute for Health and Clinical Excellence (NICE), the Canadian Agency for Drugs and Technologies in Health (CADTH), and carried out by teams of systematic reviewers in university departments of health sciences. In general, reviewers in these teams have sought to apply methods developed for systematic reviews of randomized trials to NRS. These groups include:

- Effective Practice and Organisation of Care (EPOC) Group (www.epoc.cochrane.org).
- The Centre for Reviews and Dissemination (www.york.ac.uk/inst/crd).
- EPPI centre, Institute of Education, University of London (eppi.ioe.ac.uk).
- The Effective Public Health Practice Project (EPHPP), Canadian Ministry of Health, Long-Term Care and the City of Hamilton, Public Health Services (link to list of EPHPP reviews: old.hamilton.ca/phcs/ephpp).

CRGs and Cochrane review authors have tended to limit inclusion of NRS by study design or methodological quality, acknowledging that NRS design influences susceptibility to bias. For example, the EPOC CRG accepts protocols that include interrupted time series and controlled before-and-after studies, but not other NRS designs. Other reviews have limited inclusion to studies with 'adequate methodological quality' (Taggart 2001).

13.2.3 Summary

- Review authors should carefully justify their rationale for including NRS in their systematic review.
- Review authors should consult the editorial policy of the CRG under which they propose to register their protocol concerning inclusion of NRS. Authors should consider the extent of methodological advice available in the CRG and the methodological support they have in their team.
- Review authors should specify eligibility criteria based on what researchers did (i.e. important aspects of study design), as well as factors relating to the specific review question of interest (i.e.

intervention, population, health problem), to avoid ambiguity. We suggest that authors use the items in the NRSMG checklist, or a similar checklist, to do this.

- Review authors also need information about what researchers did in primary studies to categorize the studies identified. We suggest that authors use the NRSMG lists of study design features, or a similar tool, for these purposes, and record when important aspects of study design are unclear or not reported.
- Authors reviewing questions about the adverse effects (harms) of interventions should read Chapter 14.

Table 13.2.a: List of study design features (studies with allocation to interventions at the individual level)

	RCT	Q-RCT	NRCT	CBA	PCS	RCS	HCT	NCC	CC	XS	BA	CR/CS
Was there a comparison:												
Between two or more groups of participants receiving different interventions?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N
Within the same group of participants over time?	P	P	N	Y	N	N	N	N	N	N	Y	N
Were participants allocated to groups by:												
Concealed randomization?	Y	N	N	N	N	N	N	N	N	N	na	na
Quasi-randomization?	N	Y	N	N	N	N	N	N	N	N	na	na
By other action of researchers?	N	N	Y	P	N	N	N	N	N	N	na	na
Time differences?	N	N	N	N	N	N	Y	N	N	N	na	na
Location differences?	N	N	P	P	P	P	P	na	na	na	na	na
Treatment decisions?	N	N	N	P	P	P	N	N	N	P	na	na
Participants' preferences?	N	N	N	P	P	P	N	N	N	P	na	na
On the basis of outcome?	N	N	N	N	N	N	N	Y	Y	P	na	na
Some other process? (specify)												
Which parts of the study were prospective:												
Identification of participants?	Y	Y	Y	P	Y	N	P*	Y	N	N	P	P
Assessment of baseline and allocation to intervention?	Y	Y	Y	P	Y	N	P*	Y	N	N	na	na
Assessment of outcomes?	Y	Y	Y	P	Y	P	P	Y	N	N	P	P
Generation of hypotheses?	Y	Y	Y	Y	Y	Y	Y	Y	P	P	P	na
On what variables was comparability between groups assessed:												
Potential confounders?	P	P	P	P	P	P	P	P	P	P	N	na
Baseline assessment of outcome variables?	P	P	P	Y	P	P	P	N	N	N	N	na

Y=Yes; P=Possibly; P*=Possible for one group only; N=No; na=not applicable. NB: Note that 'possibly' is used in the table to indicate cells where *either* 'Y' or 'N' may be the case. It should not be used as a response option when applying the checklist; if uncertain, the response should be 'can't tell' (see Box 13.4.a).

RCT=Randomized controlled trial; Q-RCT=Quasi-randomized controlled trial; NRCT=Non-randomized controlled trial; CBA=Controlled before-and-after study; PCS=Prospective cohort study; RCS=Retrospective cohort study; HCT=Historically controlled trial; NCC=Nested case-control study; CC=Case-control study; XS=Cross-sectional study; BA=Before-and-after comparison; CR/CS=Case report/Case series.

Table 13.2.b: List of study design features (studies with allocation to interventions at the group level)

	ClRCT	ClQ-RCT	ClNRT	CITS	CChBA	ITS	ChBA	EcoXS
Was there a comparison:								
Between two or more groups of clusters receiving different interventions?	Y	Y	Y	Y	Y	N	N	Y
Within the same group of clusters over time?	P	P	N	Y	N	Y	Y	N
Were clusters allocated to groups by:								
Concealed randomization?	Y	N	N	N	N	N	N	N
Quasi-randomization?	N	Y	N	N	N	N	N	N
By other action of researchers?	N	N	Y	P	P	N	N	N
Time differences?	N	N	N	Y	Y	Y	Y	N
Location differences?	N	N	P	P	P	N	N	P
Policy/public health decisions?	Na	na	P	P	P	P	na	na
Cluster preferences?	Na	na	P	P	P	P	na	na
Some other process? (specify)								
Which parts of the study were prospective:								
Identification of participating clusters?	Y	Y	Y	P	P	P	P	N
Assessment of baseline and allocation to intervention?	Y	Y	Y	P	P	P	P	N
Assessment of outcomes?	Y	Y	Y	P	P	P	P	N
Generation of hypotheses?	Y	Y	Y	Y	Y	Y	Y	P
On what variables was comparability between groups assessed:								
Potential confounders?	P	P	P	P	P	P	P	P
Baseline assessment of outcome variables?	P	P	P	Y	Y	Y	Y	N

Note that 'cluster' refers to an entity (e.g. an organization), not necessarily to a group of participants; 'group' refers to one or more clusters; see Box 13.4.a.

Note that 'possibly' is used in the table to indicate cells where *either* 'Y' or 'N' may be the case. It should not be used as a response option when applying the checklist; if uncertain, 'can't tell' should be used (see Box 13.4.a).

Y=Yes; P=Possibly; P*=Possible for one group only; N=No; NR=Not required. ClRCT=Cluster randomized controlled trial; ClQ-RCT=Cluster quasi-randomized controlled trial; ClNRT=Cluster non-randomized controlled trial; CITS=Controlled interrupted time series (Shadish 2002); CChBA=Controlled cohort before-and-after study (Shadish 2002); ITS=Interrupted time series; ChBA=Cohort before-and-after study (Shadish 2002); EcoXS=Ecological cross-sectional study.

13.3 Searching for non-randomized studies

13.3.1 What is different when including non-randomized studies?

13.3.1.1 Comprehensiveness of search strategy

When a review aims to include randomized trials only, a key principle of searching for eligible studies is that review authors should try as hard as possible to identify all randomized trials of the review question that have ever been started. Therefore, review authors are recommended to search trial registers, conference abstracts, grey literature, etc, as well as standard bibliographic databases such as MEDLINE, PUBMED, EMBASE (see Chapter 6). It is argued that a systematic review needs to search comprehensively in order to avoid publication biases. It is easy to argue that authors of a review that includes NRS should do the same (Petticrew 2001). However, it is important to set out the premises underpinning the original rationale for a comprehensive search and to consider very carefully whether they apply to reviews of NRS. The premises are:

a) A finite population exists of randomized trials that investigate the review question.

b) All randomized trials in this population can be identified through a search that is sufficiently comprehensive because randomized trials are relatively easily identified, registers of them are available, and they are difficult to do without funding and ethics approval, which also create an 'audit trail' (Chan 2004).

c) All randomized trials in this population, if well conducted, provide valuable information.

d) Ease of access to information about these randomized trials is related to their findings, so that the most readily identified trials may be a biased subset. This is publication bias: studies with statistically significant and favourable findings are more likely to be published in accessible places (see Chapter 10, Section 10.2). Because smaller studies are less likely to produce such findings, failure to identify all studies may result in funnel plot asymmetry. An unbiased answer can in theory be reached by identifying all randomized trials, i.e. by a comprehensive search to uncover the small, non-significant or unfavourable studies. Smaller studies may also suffer differentially from other biases, giving rise to an alternative cause of funnel plot asymmetry. The risks of these biases are reasonably well understood and may be assessed (Chapter 10, Section 10.4).

It is not clear that these premises apply equally to NRS.

Section 13.2.1.3 points out that NRS include diverse designs, and that there is difficulty in categorizing them. Even if review authors are able to set specific study design criteria against which potential NRS should be assessed for inclusion, many of the potentially eligible NRS will report insufficient information to allow them to be classified.

There is a further problem in defining exactly when a NRS comes into existence. For example, is a cohort study that has collected data on the interventions and outcome of interest, but that has not examined their association, an eligible NRS? Is computer output in a filing cabinet that includes a calculated odds ratio for the relevant association an eligible NRS? Consequently, it is difficult to define a 'finite population of NRS' for a particular review question. Some NRS that have been done may not be traceable at all, i.e. they are not to be found even in the proverbial 'bottom drawer'.

Notwithstanding the problems in defining what constitutes an eligible NRS, the actual identification of NRS provides important challenges. This is not just to do with poor reporting but also to do with:

- the absence of registers of NRS;
- poor indexing of important study design characteristics, etc;
- NRS not always requiring ethical approval (at least in the past);
- NRS not always having a research sponsor or funder; and

- NRS not always having been executed according to a pre-specified protocol.

There is no evidence that reporting biases affect randomized trials and NRS differentially. However, it is difficult to believe that reporting biases could affect NRS *less* than randomized trials, given the increasing number of features associated with carrying out and reporting randomized trials that act to prevent reporting biases which are frequently absent in NRS (pre-specified protocol, ethical approval including progress and final reports, the CONSORT statement (Moher 2001), trial registers and indexing of publication type in bibliographic databases). Unlike the situation for randomized trials, the likely magnitude and determinants of publication bias are not known.

The benefits of comprehensive searching for NRS are unclear, and this is a topic that requires further research. It is possible that the studies which are the hardest to find may be the most biased, if being hard to identify relates to poor design and small size. With reviews of randomized trials, comprehensive searching offers potential protection against bias because a defined population of eligible studies exists, so small studies with non-significant findings should, ultimately, be identified. With reviews of NRS, even if a *theoretical* finite population of eligible studies can be defined, one does not have similar confidence that missing studies with non-significant findings can be identified.

13.3.1.2 Identifying NRS in searches

It is easy to design a search strategy that identifies all evidence about an intervention by creating search strings for the population and disease characteristics, the intervention, and possibly the comparator. When a review aims to include randomized trials only, various approaches are available to restrict the search strategy to randomized trials (see Chapter 6):

a) Search for previous reviews of the review question.

b) Use resources, such as CENTRAL or CRG-specific registers, that are 'rich' in randomized trials.

c) Use methodological filters and indexing fields, such as publication type in MEDLINE, to limit searches to studies that are likely to be randomized trials.

d) Search trial registers.

To restrict the search to particular non-randomized study designs is more difficult. Of the above approaches, only (a) and (b) are likely to be at all helpful. Review authors should certainly search CRG-specific registers for potentially relevant NRS. Some CRGs (e.g. the EPOC Group) include particular types of NRS in CRG-specific registers (authors should check with their CRG). The process of identifying studies for inclusion in CENTRAL means that some, but not all, NRS are included, so searches of this database will not be comprehensive, even for studies that use a particular design. There are no databases of NRS similar to CENTRAL.

As discussed in Section 13.2.1.3, study design labels are not used consistently by authors and are not indexed reliably by bibliographic databases. Strategy (c) is unlikely to be helpful because study design labels other than randomized trial are not reliably indexed by bibliographic databases and are often used inconsistently by authors of primary studies. Some review authors have tried to develop and 'validate' search strategies for NRS (Wieland 2005, Fraser 2006, Furlan 2006). Authors have also sought to optimize search strategies for adverse effects (see Chapter 14, Section 14.5) (Golder 2006b, Golder 2006c). Because of the time consuming nature of systematic reviews that include NRS, attempts to develop search strategies for NRS have not investigated large numbers of review questions. Therefore, review authors should be cautious about assuming that previous strategies can necessarily be applied to new topics.

13.3.1.3 Reviewing citations and abstracts

Randomized trials can usually be identified in search results simply from the titles and abstracts, particularly since the implementation of reporting standards. Unfortunately, the design details of NRS that are required to assess eligibility are often not described in titles or abstracts and require access to the full study report.

13.3.2 Guidance and resources available to support review authors

The NRSMG does not recommend limiting search strategies by index terms relating to study design. However, review authors may wish to contact researchers who have reported some success in developing efficient search strategies for NRS (see Section 13.3.1) and other review authors who have carried out Cochrane reviews (or other systematic reviews) of NRS for review questions similar to their own.

When searching for NRS, review authors are recommended to search for studies investigating all effects of an intervention and not to limit search strategies to specific outcomes (Chapter 6). When searching for NRS of specific rare or long-term (usually adverse or unintended) outcomes of an intervention, including free text and MeSH terms for specific outcomes in the search strategy may be justified. Members of the Adverse Effects Methods Group have experience of doing this (see Chapter 14, Section 14.5).

Review authors should check with their CRG editors whether the CRG-specific register includes studies with particular study design features and should seek the advice of information retrieval experts within the CRG and in the Information Retrieval Methods Group (see Chapter 6, Box 6.7.a).

13.3.3 Summary

- To identify studies of the expected beneficial effects of interventions, search strategies should include search strings for the intervention and the population and health problem of interest. Currently, there are no recommended methods for restricting search strategies by study design.
- Review authors searching for evidence relating to 'suspected' adverse effects may want to consider searching for specific outcomes (i.e. adverse effects) of interest. This approach obviously cannot be used for more general searches of possible adverse effects of an intervention (see Chapter 14, Section 14.5).
- Exhaustive searching, which is recommended for randomized trials, may not be justified when reviewing NRS. However, there is no research at present to guide authors about this important issue.

13.4 Selecting studies and collecting data

13.4.1 What is different when including non-randomized studies?

Search results often contain large numbers of irrelevant citations and abstracts often do not provide adequate detail about NRS design (which are likely to be required to judge eligibility). Therefore, unlike the situation when reviewing randomized trials, very many full reports of studies may need to be obtained and read in order to select eligible studies.

Review authors need to collect all of the data required for a systematic review of randomized trials (see Chapter 7) and also data to describe (a) the features of the design of a primary study (see Section 13.2.2), (b) confounding factors considered and the methods used to control for confounding (see

Section 13.1.3), (c) aspects of risk of bias specific for NRS (see Section 13.5.1) and (d) the results (see Section 13.6.1).

Review authors normally collect 'raw' information about the results when reviewing randomized trials, e.g. for a dichotomous outcome, the total number of participants and the number experiencing the outcome in each group. If participants are randomized to groups, a comparison of these raw data is assumed to be unbiased. For a NRS, a comparison of the same raw data is 'unadjusted' and susceptible to confounding. Authors usually also report an 'adjusted' comparison estimated from a regression model which cannot be summarized in the same way. Review authors should still record the sample size recruited to each group, and the number analysed and the number of events, but also need to document any adjusted effect estimates and their standard errors or confidence intervals. These data can be used to display adjusted effect estimates and their precision in forest plots and, if appropriate, to pool data across studies.

Anecdotally, the experience of review authors is that NRS are poorly reported so that the required information is difficult to find, and different review authors may extract different information from the same paper. Data collection forms may need to be customized to the research question being investigated. Because of the diversity of potentially eligible studies and the ways in which they are reported, developing the data collection form can require several iterations in the course of reviewing a sample of primary studies. It is almost impossible to finalize these forms in advance.

Results in NRS may be presented using different measures of effect and uncertainty or statistical significance depending on the reporting style and analyses undertaken. Expert statistical advice may assist review authors to transform or 'work back' from the information provided in a paper to obtain a consistent effect measure across studies. Data collection sheets need to be able to handle the different kinds of information about study findings that authors may encounter.

13.4.2 Guidance and resources available to support review authors

As well as providing information for deciding about eligibility, the questions in Table 13.2.a and Table 13.2.b represent a convenient checklist for collecting relevant data from NRS about study design features. In using this checklist to collect information about the studies and to decide on eligibility, the intention should be to document what researchers did in the primary studies, rather than what researchers called their studies or think they did. Items should be recorded as 'Yes', 'No' or 'Can't tell'. Box 13.4.a provides guidance on using these tables as checklists.

Data collection forms have been developed for use in NRSMG workshops to illustrate data extraction from NRS. These include: the study design checklist, templates for collecting information about confounding factors, their comparability at baseline, methods used to adjust for confounding, and effect estimates. These resources (available from the *Handbook* resource web site, www.cochrane.org/resources/handbook) can be used as a guide to the types of data collection forms that review authors will need. However, review authors will need to customize the forms carefully for the review question being studied.

Box 13.4.a: User guide for data collection/study assessment using checklist in Table 13.2.a or Table 13.2.a

Note: Users need to be very clear about the way in which the terms 'group' and 'cluster' are used in these tables. Table 13.2.a only refers to groups, which is used in its conventional sense to mean a number of individual participants. With the exception of allocation on the basis of outcome, 'group' can be interpreted synonymously with 'intervention group'. Table 13.2.b refers to both clusters and

groups. In this table, 'clusters' are typically an organizational entity such as a family health practice, or administrative area, not an individual. As in Table 13.2.a, 'group' is synonymous with 'intervention group' and is used to describe a collection of allocated units, but in Table 13.2.b these units are clusters rather than individuals. Furthermore, although individuals are nested in clusters, a cluster does not necessarily represent a fixed collection of individuals. For instance, in cluster-allocated studies, clusters are often studied at two or more time-points (periods) with different collections of individuals contributing to the data collected at each time-point.

Was there a comparison?

Typically, researchers compare two or more groups that receive different interventions; the groups may be studied over the same time period, or over different time periods (see below). Sometimes researchers compare outcomes in just one group but at two time-points. It is also possible that researchers may have done both, i.e. studying two or more groups and measuring outcomes at more than one time-point.

Were participants/clusters allocated to groups by?

These items aim to describe how groups were formed. None will apply if the study does not compare two or more groups of subjects. The information is often not reported or is difficult to find in a paper. The items provided cover the main ways in which groups may be formed. More than one option may apply to a single study, although some options are mutually exclusive (i.e. a study is either randomized or not).

Randomization: Allocation was carried out on the basis of truly random sequence. Such studies are covered by the standard guidance elsewhere in this *Handbook*. Check carefully whether allocation was adequately concealed until subjects were definitively recruited.

Quasi-randomization: Allocation was done on the basis of a pseudo-random sequence, e.g. odd/even hospital number or date of birth, alternation. Note: when such methods are used, the problem is that allocation is rarely concealed. These studies are often included in systematic reviews that only include randomized trials, using assessment of the risk of bias to distinguish them from properly randomized trials.

By other action of researchers: This is a catch-all category and further details should be noted if the researchers report them. Allocation happened as the result of some decision or system applied by the researchers. For example, subjects managed in particular 'units' of provision (e.g. wards, general practices) were 'chosen' to receive the intervention and subjects managed in other units to receive the control intervention.

Time differences: Recruitment to groups did not occur contemporaneously. For example, in a historically controlled study subjects in the control group are typically recruited earlier in time than subjects in the intervention group; the intervention is then introduced and subjects receiving the intervention are recruited. Both groups are usually recruited in the same setting. If the design was under the control of the researchers, both this option and 'other action of researchers' must be ticked for a single study. If the design 'came about' by the introduction of a new intervention, both this option and 'treatment decisions' must be ticked for a single study.

Location differences: Two or more groups in different geographic areas were compared, and the choice of which area(s) received the intervention and control interventions was not made randomly. So, both this option and 'other action of researchers' could be ticked for a single study.

Treatment decisions: Intervention and control groups were formed by naturally occurring variation in treatment decisions. This option is intended to reflect treatment decisions taken mainly by the clinicians responsible; the following option is intended to reflect treatment decisions made mainly on the basis of subjects' preferences. If treatment preferences are uniform for particular provider 'units', or switch over time, both this option and 'location' or 'time' differences should be ticked.

Patient preferences: Intervention and control groups were formed by naturally occurring variation in patients' preferences. This option is intended to reflect treatment decisions made mainly on the basis of subjects' preferences; the previous option is intended to reflect treatment decisions taken mainly by the clinicians responsible.

On the basis of outcome: A group of people who experienced a particular outcome of interest were compared with a group of people who did not, i.e. a case-control study. Note: this option should be ticked for papers that report analyses of *multiple risk factors for a particular outcome* in a large series of subjects, i.e. in which the total study population is divided into those who experienced the outcome and those who did not. These studies are much closer to nested case-control studies than

cohort studies, even when longitudinal data are collected prospectively for consecutive patients.

Additional options for cluster-allocated studies

Location differences: see above.

Policy/public health decisions: Intervention and control groups were formed by decisions made by people with the responsibility for implementing policies about public health or service provision. Where such decisions are coincident with clusters, or where such people are the researchers themselves, this item overlaps with 'other action of researchers' and 'cluster preferences'.

Cluster preferences: Intervention and control groups were formed by naturally occurring variation in the preferences of clusters, e.g. preferences made collectively or individually at the level of the cluster entity.

Which parts of the study were prospective?

These items aim to describe which parts of the study were conducted prospectively. In a randomized controlled trial, all four of these items would be prospective. For NRS it is also possible that all four are prospective, although inadequate detail may be presented to discern this, particularly for generation of hypotheses. In some cohort studies, participants may be identified, and have been allocated to treatment retrospectively, but outcomes are ascertained prospectively.

On what variables was comparability of groups assessed?

These questions should identify 'before-and-after' studies. Baseline assessment of outcome variables is particularly useful when outcomes are measured on continuous scales, e.g. health status or quality of life.

Response options

Try to use only 'Yes', 'No' and 'Can't tell' response options. 'N/a' should be used if a study does not report a comparison between groups.

13.4.3 Summary

- Reviewing citations and abstracts identified by searching will be very time consuming, first because of the volume of citations identified and second because the information needed to judge eligibility may not be reported in the title or abstract.
- Collect data as for a randomized trial (i.e. details of study, study population, sample size recruited, sample size analysed, etc).
- Collect data about what researchers did (NRSMG checklist, or similar).
- Collect data about the confounding factors considered.
- Collect data about the comparability of groups on confounding factors considered.
- Collect data about the methods used to control for confounding.
- Collect data about multiple effect estimates (both unadjusted and adjusted estimates, if available).

13.5 Assessing risk of bias in non-randomized studies

13.5.1 What is different when including non-randomized studies?

13.5.1.1 Sources of bias in non-randomized studies

Bias may be present in findings from NRS in many of the same ways as in poorly designed or conducted randomized trials (see Chapter 8). For example, numbers of exclusions in NRS are frequently unclear, intervention and outcome assessment are often not conducted according to standardized protocols, and outcomes may not be assessed blind. The biases caused by these problems are likely to be similar to those that occur in randomized trials, and review authors should be familiar

with Chapter 8 that describes these issues. None of these problems are any less difficult to overcome in a well-planned non-randomized prospective study than in a randomized trial.

In NRS, use of allocation mechanisms other than concealed randomization means that groups are unlikely to be comparable. These potential systematic differences between characteristics of participants in different intervention 'groups' are likely to be the issue of key concern in most NRS, and we refer to this as selection bias. When selection bias produces imbalances in prognostic factors associated with the outcome of interest then 'confounding' is said to occur. Statistical methods are sometimes used to counter bias introduced from confounding by producing 'adjusted' estimates of intervention effects, and part of the assessment of study quality may involve making judgements about the appropriateness of the analysis as well as the design and execution of the study.

The variety of study designs classified as NRS, and their varying susceptibility to different biases, makes it difficult to produce a generic robust tool that can be used to evaluate risk of bias. Within a review that includes NRS of different designs, several tools for assessment of risk of bias may need to be created. Inclusion of a knowledgeable methodologist in the review team is essential to identify the key areas of weakness in the included study designs.

With randomized trials, assessment of the risk of bias focuses on systematic bias, which is usually assumed to be 'optimistic' in direction. The tendency for researchers to design, execute, analyse and report their primary studies to give the findings that are expected, consciously or subconsciously, is also likely to apply to NRS where researchers have control over key decisions (e.g. allocation to intervention, or selection of centres). In truly observational NRS, bias arising from 'confounding by indication' may not be so consistent; healthcare professionals may have differing opinions about the appropriateness of alternative interventions for their patients, contingent on the patients' presenting severity of illness or co-morbidities. Differences in case-mix between locations that are being compared may be haphazard. Therefore, when reviewing NRS, the variability of biases and the between-study heterogeneity they induce is at least as important as systematic bias when reviewing NRS.

13.5.1.2 Evidence of risk of bias in non-randomized studies

Some insight into the risk of bias in non-randomized studies can be obtained by comparing randomized trials at low risk of bias with randomized trials at high risk of bias. Controlled trials that allocate participants by quasi-randomization, or that fail to conceal allocation during recruitment, are at risk of selection bias, just like a prospectively conducted, overtly non-randomized, trial or cohort study. Chapter 8 reviews evidence on several aspects of risk of bias in randomized trials, and points out that methodological limitations in randomized trials tend to exaggerate the beneficial effects of interventions.

Researchers have also compared the findings of separate meta-analyses of randomized trials and NRS of the same research question, assuming that such methodological systematic reviews provide a way to investigate the risk of bias in NRS. Some reviews of this kind have reported discrepancies by study design but fair comparisons are very difficult to make (MacLehose 2000). There are at least two reasons for this:

- Randomized trials and NRS of precisely the same question are rare; for example, studies of the same intervention using different study designs usually differ systematically with respect to the population, intervention or outcome.
- Randomized trials and NRS may differ systematically in several ways with respect to their risk of bias (reporting biases as well as selection, performance, detection and attrition biases), and NRS are frequently of relatively poor quality.

These reasons may explain the inconsistent conclusions from methodological systematic reviews that have compared findings from randomized trials and NRS of the same research question. Deeks et al. reviewed eight such reviews (Deeks 2003), and found that:

- 5/8 concluded that there were differences between effects estimated by randomized trials and NRS for many but not all interventions, with no consistent pattern;
- 1/8 concluded that NRS overestimated the effect [benefit] for all interventions studied;
- 2/8 concluded that the effects estimated by randomized trials and NRS were "remarkably similar".

A similar methodological review compared the findings of randomized trials and patient preference studies (King 2005). The review concluded that there is little evidence that preferences "significantly affect validity", such that preferences did not appear to confound intervention effects.

Some considerations in the interpretation of these sorts of empirical studies are relevant. First, both the publication of primary studies and the selection of primary studies by review authors may be biased. There is also the possibility of bias in their classification of the review findings. Deeks et al. found that the same comparison was sometimes classified as discrepant in one review and comparable in a second. This highlights the difficulty of defining what represents a 'difference'.

Second, the observation that differences were not consistently optimistic remains an important one and is consistent with the principle that effect estimates from NRS are more heterogeneous than expected by chance (Greenland 2004). Some empirical evidence for this comes from innovative simulation studies (Deeks 2003). Deeks et al. pointed out that biases in NRS are highly variable, and may best be considered as introducing extra uncertainty in the results rather than an estimable systematic bias. This uncertainty acts over and above that accounted for in confidence intervals, and in large studies may easily be 5 to 10 times the magnitude of the 95% confidence interval.

Finally, methodological reviews are caught in a circular loop: they need to assume either that NRS are valid and hence differences between effect estimates from randomized trials and NRS are also valid and can be attributed to external factors, or that NRS are biased and hence differences between effect estimates from randomized trials and NRS can be explained by differential risk of bias. The truth may well lie somewhere in between these extremes, but the fact remains that methodological reviews cannot unequivocally partition discrepancies to different sources. Moreover, if multiple factors distinguish randomized trials and NRS and influence effect size, then observing no difference between the effect sizes estimated from randomized trials and NRS can also be explained as the consequence of effects of multiple factors influencing the effect of an intervention in different directions. It is not logical to assume that finding no difference means that NRS are valid and finding a difference means that NRS are not valid.

13.5.2 Guidance and resources available to support review authors

13.5.2.1 General considerations in assessing risk of bias in non-randomized studies

Reporting of randomized trials is relatively straightforward and, increasingly, guided by the CONSORT statement (Moher 2001). A similar consensus statement, STROBE, for the reporting of observational epidemiological studies has been developed, although much more recently (Vandenbroucke 2007, von Elm 2007). Therefore, the quality of reporting of information required to assess the risk of bias is likely to be less good for NRS. This is likely to hinder any assessment of risk of bias.

A protocol is a tool to protect against bias; when registered in advance of a study starting, it proves that aspects of study design and analysis were considered in advance of starting to recruit, and that data definitions and methods for standardizing data collection were defined. Because of the need for research ethics approval, all randomized trials must have a protocol, even if protocols vary in their quality and the items that they specify; many randomized trials, particularly those sponsored by industry, also have detailed study manuals. Historically, researchers have not had to obtain research ethics approval for many NRS, and primary NRS rarely report whether the methods are based on a protocol. Therefore, the protection offered by a protocol often does not exist for NRS. The implications of not having a protocol have not been researched. However, it means, for example, that there is no constraint on the tendency of researchers to 'cherry-pick' outcomes, subgroups and analyses to report, which happens to a greater or lesser extent even in randomized trials where protocols exist (Chan 2004).

In common with randomized trials, dimensions of bias to be assessed include selection bias (concerning comparability of groups, confounding and adjustment), performance bias (concerning the fidelity of the interventions, and quality of the information regarding who received what interventions, including blinding of participants and healthcare providers), detection bias (concerning unbiased and correct assessment of outcome, including blinding of assessors), attrition bias (concerning completeness of sample, follow-up and data) and reporting bias (concerning publication biases and selective reporting of results). Assessment of risk of bias in randomized trials has developed by identifying the design features which are used to prevent each of these dimensions, and noting whether each trial fulfils the requirements. Risk of bias assessments for NRS should proceed in the same way, with pre-specification of the features to be assessed in the protocol, recording what happened in the study, and a judgement of whether this was adequate, inadequate or unclear as a method to avoid risk of this particular bias. Determining these features is likely to require expert input from an epidemiologist, and will depend in part on the clinical question. Particular care should be given to the assessment of confounding (see Section 13.5.2.2).

The reason for careful attention to the design *features* of primary studies (such as how participants were allocated to groups, or which parts of the study were prospective) rather than design *labels* (such as 'cohort' or 'cross-sectional') is because it is hypothesized that the risk of bias is influenced by the specific features of a study rather than a broad categorization of the approach taken. Furthermore, terms such as 'cohort' and 'cross-sectional' are ambiguous and cover a diverse range of specific study designs. No empirically-derived list is available of study design features that are relevant to the risk of bias, although a shortlist can be constructed from evidence and theory about the risk of bias in aetiological studies and randomized trials (see Section 13.2.2 and 13.4.2).

Because of the diversity of NRS, different methods may be needed to assess NRS with different design features. One important distinction is between studies in which allocation to groups is by outcome (e.g. case-control studies) and studies in which allocation to groups is more directly related to interventions. In the former type of study, it is the exposure of interest, rather than the outcome, that is most susceptible to bias; review authors need to ask whether researchers assessing the exposure were masked to whether participants had experienced the outcome or not (i.e. were cases or controls). Case-control studies are well suited to investigating associations between rare outcomes and multiple exposures, so may have an important role in generating evidence about the potential adverse effects and unintended beneficial effects of interventions. They have also been used to evaluate large-scale public health interventions such as accident prevention and screening (MacLehose 2000), which are difficult or expensive to evaluate by randomized trials. However, review authors should familiarize themselves with epidemiological considerations that particularly apply to such studies (Rothman 1986). Note that some analyses of patient registries also have similarities with case-control studies: for example, if the entire database is divided into groups of patients who have or have not experienced a particular outcome and exposures associated with the outcome are investigated. Review authors

require a deeper knowledge of epidemiology when assessing the risk of bias in NRS, compared with randomized trials.

13.5.2.2 Confounding and adjustment

Researchers do not always make the same decisions concerning confounding factors, so the method used to control for confounding is an important source of heterogeneity between studies. There may be differences in the confounding factors considered, the method used to control for confounding and the precise way in which confounding factors were measured and included in analyses. Many (but not all) NRS describe the confounding factors that were considered and whether confounding was taken into account by the study design or analysis; most also report the baseline characteristics of the groups being compared. However, assessing what researchers actually did to control for confounding may be difficult; far fewer studies describe precisely how confounding factors were measured or fitted as covariates in regression models (e.g. as a continuous, ordinal, or grouped categorical variable).

Some specific suggestions for assessing risk of selection bias are as follows.

- At the stage of writing the protocol, list potential confounding factors.
- Identify the confounding factors that the researchers have considered and those that have been omitted. Note the ways in which they have been measured (the ability to control for a confounding factor depends on the precision with which the factor is measured).
- Assess the balance between comparator groups at baseline with respect to the main prognostic or confounding factors.
- Identify what researchers did to control for selection bias, i.e. any design features used for this purpose (e.g. matching or restriction to particular subgroups) and the methods of analysis (e.g. stratification or regression modelling with propensity scores or covariates).

There is no established method for identifying a pre-specified set of important confounders. Listing potential confounding factors should certainly be done 'independently' and, one might argue, 'systematically'. The list should not be generated solely on the basis of factors considered in primary studies included in the review (at least, not without some form of independent validation), since the number of potential confounders is likely to increase over time (hence, older studies may be out of date) and researchers themselves may simply choose to measure confounders considered in previous studies (hence, such a list could be selective). (Researchers investigating aetiological associations often do not explain their choice of confounding factors (Pocock 2004).) Rather, the list should be based on evidence (although undertaking a systematic review to identify all potential prognostic factors is extreme) and expert opinion from members of the review team and advisors.

Reporting results of assessments of confounders in a Cochrane review may best be achieved by creating additional tables listing the pre-stated confounders as columns, the studies as rows, and indicating whether each study: (i) restricted participant selection so that all groups had the same value for the confounder (e.g. restricting the study to male participants only); (ii) demonstrated balance between groups for the confounder; (iii) matched on the confounder; or (iv) adjusted for the confounder in statistical analyses to quantify the effect size.

13.5.2.3 Tools for assessing methodological quality or risk of bias in non-randomized studies

Chapter 8 (Section 8.5) describes the 'Risk of bias' tool that review authors are expected to use for assessing risk of bias in randomized trials. This involves consideration of six features: sequence generation, allocation sequence concealment, blinding, incomplete outcome data, selective outcome reporting and 'other' potential sources of bias. Items are assessed by: (i) providing a description of

what happened in the study; (ii) providing a judgement on the adequacy of the study with regard to the item. The judgement is formulated by answering a pre-specified question, such that an answer of 'Yes' indicates low risk of bias, an answer of 'No' indicates high risk of bias, and an answer of 'Unclear' indicates unclear or unknown risk of bias. The tool was not developed with NRS in mind, and the six domains are not necessarily appropriate for NRS. However, the general structure of the tool and the assessments seems useful to follow when creating risk of bias assessments for NRS.

For experimental and controlled studies, and for prospective cohort studies (see Box 13.1.a and Section 13.2.2), the six domains in the standard 'Risk of bias' tool could usefully be assessed, whether allocation is randomized or not. This is the minimum assessment review authors should carry out and more details will usually be required. An additional component is to assess the risk of bias due to confounding. The depth of this assessment is likely to depend on the heterogeneity between studies and whether the review authors propose a quantitative synthesis (see Section 13.6). If studies are heterogeneous and no quantitative synthesis is proposed, then a less detailed assessment can nevertheless serve the purposes of illustrating the heterogeneity and informing interpretation of the findings of the review.

Many instruments for assessing methodological quality of non-randomized studies of interventions have been created, and were reviewed systematically by Deeks et al. (Deeks 2003). In their review they located 182 tools, which they reduced to a shortlist of 14, and identified six as potentially useful for systematic reviews as they "force the reviewer to be systematic in their study assessments and attempt to ensure that quality judgements are made in the most objective manner possible". However, all six required a degree of adjustment as they neglected to elicit detailed information about how study participants were allocated to groups, which in terms of the risk of selection bias is likely to be critical. Not all of the six tools were suitable for different study designs. In common with some tools for assessing the quality of randomized trials, some did not distinguish items relating to the quality of the study and the quality of reporting of the study. The two most useful tools identified in this review are the Downs and Black instrument and the Newcastle-Ottawa Scale (Downs 1998, Wells 2008).

The Downs and Black instrument has been modified for use in a methodological systematic review (MacLehose 2000). The reviewers found that some of the 29 items were difficult to apply to case-control studies, that the instrument required considerable epidemiological expertise and that it was time consuming to use. The Newcastle-Ottawa Scale, which has been used in NRSMG workshops to illustrate issues in data extraction from primary NRS, contains only eight items and is simpler to apply (Wells 2008). However, the items may still need to be customized to the review question of interest. Review authors also need to be aware of differences in epidemiological terminology in different countries; for example, the Newcastle-Ottawa Scale uses the term 'selection bias' to describe what others may call 'applicability' or 'generalizability'.

Acknowledging the importance of distinguishing between 'what researchers do' and 'what researchers report', review authors may also find it helpful to consider items included in reporting statements for randomized trials (Moher 2001) and observational epidemiological studies (Vandenbroucke 2007) in order to highlight gaps in reporting (and execution) in NRS (Reeves 2004, Reeves 2007).

13.5.2.4 Practical limitations in assessing risk of bias in non-randomized studies

Two studies of systematic reviews that included NRS have commented that only a minority of reviews assessed the methodological quality of included studies (Audige 2004, Golder 2006a). Members of the NRSMG have gained experience of trying to assess risk of bias in non-randomized studies. Anecdotally, review authors have reported that NRS are generally of poor methodological quality, or are poorly reported so that assessing methodological quality and risk of bias consistently across primary studies is difficult or impossible (Kwan 2004). Even the Newcastle-Ottawa scale has been

reported to be difficult to apply, so agreement between review authors is likely to be modest. Methodological information can be difficult to find in papers, making the task frustrating, especially when using some of the more detailed instruments; review authors may spend a long time searching for details of what researchers did, only to conclude that the information was not reported. Nevertheless, collecting some factual information (for example, the confounders considered and what researchers did about confounding) can still be useful since such information illustrates the extent of heterogeneity between studies.

13.5.3 Summary

- At the stage of writing the protocol for the review, compile a list of potential confounding factors and justify the choice.
- At the stage of writing the protocol for the review, decide how the risk of bias in primary studies will be assessed, including the extent of control for confounding.
- For NRS conducted entirely prospectively, apply the methods that the Collaboration recommends for randomized trials.
- There is no single recommended instrument, so review authors are likely to need to include supplementary risk of bias instruments or items.
- Issues such as confounding cannot easily be addressed with in the format of the new risk of bias tool and require creation of additional tables for reporting assessments.
- Collecting some factual information (for example, the confounders considered and what researchers did about confounding) is useful since such information illustrates the extent of heterogeneity between studies.
- Review authors who choose to include case-control studies in a Cochrane review should ensure that they are familiar with common pitfalls that can affect such studies and that they assess their susceptibility to bias using an instrument designed for this purpose.
- Review authors may decide that collecting great detail about the risk of confounding and other biases is not warranted. However, if this approach is taken, review authors must acknowledge the potential extent of the heterogeneity between studies with respect to potential residual confounding and other biases and demonstrate that they have considered this source of heterogeneity in their interpretation of the findings of the primary NRS reviewed.

13.6 Synthesis of data from non-randomized studies

13.6.1 What is different when including non-randomized studies?

Review authors should expect greater heterogeneity in a systematic review of NRS than a systematic review of randomized trials. This is due to the increased potential for methodological diversity through variation between primary studies in their risk of selection bias, variation in the way in which confounding is considered in the analysis and greater risk of other biases through poor design and execution. There is no way of controlling for these biases in the analysis of primary studies and no established method for assessing how, or the extent to which, these biases affect primary studies (but see Chapter 8).

There is a body of opinion that it is appropriate to pool results of non-randomized studies when they have large effects, but the logic of this view can be questioned. NRS with large effects are as likely (perhaps more likely) to be biased and to be heterogeneous as NRS with small effects. Judgements about the risk of bias and heterogeneity should be based on critical appraisal of the characteristics and methods of included studies, not on their results.

When assessing similarity of studies prior to a meta-analysis, review authors should also keep in mind that some features of studies, for example assessment of outcome not masked to intervention allocation, may be relatively homogeneous across NRS but still leave all studies at risk of bias.

If authors judge that included NRS are both reasonably resistant to biases and relatively homogeneous in this respect, they may wish to combine data across studies using meta-analysis (Taggart 2001). Unlike for randomized trials, it will usually be appropriate to analyse adjusted, rather than unadjusted, effect estimates, i.e. analyses that attempt to 'control for confounding'. This may require authors to choose between alternative adjusted estimates reported for one study. Meta-analysis of adjusted estimates can be performed as an inverse-variance weighted average, for example using the 'Generic inverse-variance' outcome type in RevMan (see Chapter 9, Section 9.4.3). In principle, any effect measure used in meta-analysis of randomized trials can also be used in meta-analysis of non-randomized studies (see Chapter 9, Section 9.2), although the odds ratio will commonly be used as it is the only effect measure for dichotomous outcomes that can be estimated from case-control studies, and is estimated when logistic regression is used to adjust for confounders.

One danger is that a very large NRS of poor methodological quality (for example based on routinely collected data) may dominate the findings of other smaller studies at less risk of bias (perhaps carried out using customized data collection). Authors need to remember that the confidence intervals for effect estimates from larger NRS are less likely to represent the true uncertainty of the observed effect than are the confidence intervals for smaller NRS (see Section 13.5.1.2), although there is no way of estimating or correcting for this.

13.6.2 Guidance and resources available to support review authors

13.6.2.1 Controlling for confounding

Imbalances in prognostic factors in NRS (e.g. 'confounding by indication' (Grobbee 1997)) must be accounted for in the statistical analysis. There are several methods to control for confounding. Matching, i.e. the generation of similar intervention groups with respect to important prognostic factors, can be used to lessen confounding at the study design stage. Stratification and regression modelling are statistical approaches to control for confounding, which result in an estimated intervention effect adjusted for imbalances in observed prognostic factors. Some analyses use propensity score methods as part of a two-stage analysis. The probability of an individual receiving the experimental intervention (the propensity score) is first estimated according to their characteristics using a logistic regression model. This single summary measure of case-mix is then used for matching, stratification or in a regression model.

Matching

The selection of patients with similar values for important prognostic factors results in more comparable groups. Therefore, matching can be seen as a type of confounder adjustment. Matching can be either at the level of individual patients (i.e. one or more control participants are selected who have similar characteristics to an intervention participants) or at the level of participants strata (i.e. selecting participants so that there are roughly the same number of control participants in one stratum, for example 60 years or older, as in the intervention group). Where direct matching has been used, the paired nature of the data has to be considered in the statistical analysis of a single study in order to obtain appropriate confidence intervals for the estimated effect of the intervention. Matching on a single measure such as the propensity score is easier to achieve than matching individuals with a particular set of characteristics.

Stratification

Stratification involves the division of participants into subgroups with respect to categorical (or categorized quantitative) prognostic factors, for example classifying age into decades, or weight into quartiles. The intervention effect is then estimated in each stratum and a pooled estimate is calculated across strata. This procedure can be interpreted as a meta-analysis at the level of an individual study. For dichotomous outcomes, the Mantel-Haenszel method is often used to estimate the overall intervention effect, with versions available for the odds ratio, the risk ratio and the risk difference as measures of intervention effect. Again, the propensity score may be used as the stratification variable.

Modelling

In a modelling approach, information on intervention and prognostic factors is incorporated into a regression equation. Advantages of regression models include the possibility of incorporating quantitative factors without categorization and the possibility of modelling trends in confounders measured on an ordinal scale. For dichotomous outcomes, a logistic regression model is almost always used to estimate the adjusted intervention effect. Thus, the odds ratio is (implicitly) used as the measure of intervention effect. Regression models are also available for risk ratio and absolute risk reduction measures of effect but these models are rarely used in practice. A linear regression model is typically used for continuous outcomes (perhaps after transformation of one or more variables), and a proportional hazards regression (Cox regression) model is typically used for time-to-event data. Regression models may also use the propensity score alone or in combination with other participant characteristics as explanatory variables.

Review authors should acknowledge that in any non-randomized study, even when experimental and control groups appear comparable at baseline, the effect size estimate is still at risk of bias due to residual confounding. This is because all methods to control for confounding are imperfect, for example for the following reasons.

- Unknown, and consequently unmeasured, confounding factors, which cannot be controlled for.
- Poor resolution in the measurement of confounders, e.g. co-morbidity assessed on a simple ordinal scale (Concato 1992), which represents non-differential error misclassification with respect to confounders.
- Practical constraints on the resolution of matching, and the number of confounders on which participants can be matched, in matched analyses.
- Poor resolution in the way confounders are measured in stratified analyses, or handled in analyses, illustrated by the width of strata (e.g. decades of age); this limitation also applies to regression models when confounders are categorized and modelled discretely.
- Assumptions in the way confounders are modelled in regression analyses, because of imperfect knowledge of the shape of the association between confounder and outcome.

There is no established method for judging the likely extent of residual confounding. The direction of bias from confounding is unpredictable and may differ between studies.

13.6.2.2 Combining studies

Estimated intervention effects for different study designs can be expected to be influenced to varying degrees by different sources of bias (see Section 13.5). Results from different study designs should be expected to differ systematically, resulting in increased heterogeneity. Therefore, we recommend that NRS which used different study designs (or which have different design features), or randomized trials and NRS, should not be combined in a meta-analysis.

Because of the need to control for confounding as best as possible, the estimated intervention effect and its standard error (or confidence interval) are key pieces of information which should be used for pooling NRS in a meta-analysis. (Simple numerators and denominators, or means and standard errors, for intervention and control groups cannot control for confounding unless the groups have been matched at the design stage.) Consequently, meta-analysis methods based on estimates and standard errors, and in particular the generic inverse-variance method, will be suitable for NRS (see Chapter 9, Section 9.4.3).

It is straightforward to extract an adjusted effect estimate and its standard error for a meta-analysis if a single adjusted estimate is reported for a particular outcome in a primary NRS. However, many NRS report both unadjusted and adjusted effect estimates, and some NRS report multiple adjusted estimates from analyses including different sets of covariates. Review authors should record both unadjusted and adjusted effect estimates but it can be difficult to choose between alternative adjusted estimates. No general recommendation can be made for the selection of which adjusted estimate is preferable. Possible selection rules are:

- use the estimate from the model that adjusted for the maximum number of covariates;
- use the estimate that is identified as the primary adjusted model by the authors; and
- use the estimate from the model that includes the largest number of confounders considered important at the outset by the review authors.

Sensitivity analyses could be performed by pooling separately the most optimistic and pessimistic results from each included study.

There is a subtle statistical point regarding the different interpretation of adjusted and unadjusted effects when expressed as odds or hazard ratios. The unadjusted effect estimate is known as the population average effect, and if the estimate were unbiased would be the effect of intervention observed in a population with an average mixture of prognostic characteristics. When estimates are adjusted for prognostic characteristics, the estimated effects are known as conditional estimates and are the intervention effects that would be observed in groups with particular combinations of the adjusted covariates. Mathematical research has shown that conditional estimates are usually larger (further from an OR or HR of 1) than population average estimates. This phenomenon may not be observed in systematic reviews due to heterogeneity in the estimates of the studies.

13.6.2.3 Analysis of heterogeneity

The exploration of possible sources of heterogeneity between studies should be part of any Cochrane review, and is discussed in detail in Chapter 9 (Section 9.6). Non-randomized studies may be expected to be more heterogeneous than randomized trials, given the extra sources of methodological diversity and bias. The simplest way to show the variation in results of studies is by drawing a forest plot (see Chapter 11, Section 11.3.2).

It may be of value to undertake meta-regression analyses to identify important determinants of heterogeneity, even in reviews where studies are considered too heterogeneous to pool. Such analyses may help to identify methodological features which systematically relate to observed intervention effects, and help to identify the subgroups of studies most likely to yield valid estimates of intervention effects.

13.6.2.4 When pooling is judged not to be appropriate

Before undertaking a meta-analysis, review authors must ask themselves the standard question about whether primary studies are ‘similar enough’ to justify pooling (see Chapter 9). Forest plots in RevMan allow the presentation of estimates and standard errors for each study, using the ‘Generic

inverse-variance' outcome type. Meta-analyses can be suppressed, or included only for subgroups within a plot. Providing that effect estimates from the included studies can be expressed using consistent effect measures, we recommend that review authors display individual study results for NRS with similar study design features using forest plots, as a standard feature. If consistent effect measures are not available, then additional tables should be used to present results in a systematic format.

If included studies are not sufficiently homogeneous to combine in a meta-analysis (which is expected to be the norm for reviews that include NRS), the NRSMG recommends displaying the results of included studies in a forest plot but suppressing the pooled estimate. Studies may be sorted in the forest plot (or shown in separate forest plots) by study design feature, or some other feature believed to reflect susceptibility to bias (e.g. number of Newcastle-Ottawa Scale 'stars' (Wells 2008)). Heterogeneity diagnostics and investigations (e.g. a test for heterogeneity, the I^2 statistic and meta-regression analyses) are worthwhile even when a judgement has been made that calculating a pooled estimate of effect is not (Higgins 2003, Siegfried 2003).

Narrative syntheses are, however, problematic, because it is difficult to set out or describe results without being selective or emphasizing some findings over others. Ideally, authors should set out in the review protocol how they plan to use narrative synthesis to report the findings of primary studies.

13.6.3 Summary

- Heterogeneity will be greater in a systematic review of NRS than in a systematic review of randomized trials. Therefore, authors should consider very carefully the likely extent of heterogeneity between included studies when deciding whether to pool findings quantitatively (i.e. by meta-analysis). We expect pooling of effect estimates from NRS to be the exception, rather than the rule.
- Effect estimates from NRS should not be combined with effect estimates from randomized trials, or across NRS that have dissimilar study design features.
- Forest plots should be used to summarize the findings from included studies.
- Heterogeneity diagnostics and investigations may be used irrespective of whether or not a decision has been taken to pool effect estimates from different studies.

13.7 Interpretation and discussion

13.7.1 Challenges in interpreting Cochrane reviews of effectiveness that include non-randomized studies

Review authors face great challenges in demonstrating convincingly that the result of a Cochrane review of NRS can give anything close to a definitive answer about the likely effect of an intervention (Deeks 2003). In many situations, reviews of NRS are likely to conclude that calculating an 'average' effect is not helpful (Siegfried 2003), that evidence from NRS is inadequate to prove effectiveness or harm (Kwan 2004) and that randomized trials should be undertaken (Taggart 2001).

Challenges arise at all stages of conducting a review of NRS: deciding which study designs to include, searching for studies, assessing studies for potential bias, and deciding whether to pool results. A review author needs to satisfy the reader of the review that these challenges have been adequately addressed, or should discuss how and why they cannot be met. In this section, the challenges are

illustrated with reference to issues raised in the different sections of this chapter. The Discussion section of the review should address the extent to which the challenges have been met.

13.7.1.1 Have all important and relevant studies been included?

Even if the choice of eligible study designs can be justified, it may be difficult to show that all relevant studies have been identified because of poor indexing and inconsistent use of study design labels by researchers. Comprehensive search strategies that focus only on the health condition and intervention of interest are likely to result in a very long list of citations including relatively few eligible studies; conversely, restrictive strategies will inevitably miss some eligible studies. In practice, available resources may make it impossible to process the results from a comprehensive search, especially since authors will often have to read full papers rather than abstracts to determine eligibility. The implications of using a more or less comprehensive search strategy are not known.

13.7.1.2 Has the risk of bias to included studies been adequately assessed?

Interpretation of the results of a review of NRS must include consideration of the likely direction and magnitude of bias. Biases that affect randomized trials also affect NRS but typically to a greater extent. For example, attrition in NRS is often worse (and poorly reported), intervention and outcome assessment are rarely conducted according to standardized protocols, and outcomes are rarely blind. Too often these limitations of NRS are seen as part of doing a NRS, and their implications for risk of bias are not properly considered. For example, some users of evidence may consider NRS that investigate long-term outcomes to have 'better quality' than randomized trials of short-term outcomes, simply on the basis of their relevance without appraising their risk of bias (see Section 13.2.1.4).

Assessing the magnitude of confounding in NRS is especially problematic. Review authors must not only have adequate methods for assessment but also collect and report adequate detail about the confounding factors considered by researchers and the methods used to control for confounding. The information may not be available from the reports of the primary studies, preventing the review authors from investigating differences in the methods of eligible studies and other sources of heterogeneity that were considered likely to be important when the protocol was written.

Authors must remember the following points about confounding:

- The direction of the bias introduced by confounding is unpredictable.
- Methods used by researchers to control for confounding are like to vary between studies.
- The extent of residual confounding in any particular study is unknown, and is likely to vary between studies.
- Residual confounding (and other biases) means that confidence intervals underestimate the true uncertainty around an effect estimate.
- It is important to identify the likely confounding factors that have not been adjusted for, as well as those that have been adjusted for.

The challenges described above affect all systematic reviews of NRS. However, challenges may be less extreme in some healthcare areas (e.g. confounding may be less of a problem in observational studies of long-term or adverse effects, or some public health primary prevention interventions).

One clue to the presence of bias is notable between-study heterogeneity. Although heterogeneity can arise through differences in participants, interventions and outcome assessments, the possibility that bias is the cause of heterogeneity in reviews of NRS must be considered seriously. However, lack of heterogeneity does not indicate lack of bias, since it is possible that a consistent bias applies in all studies.

Can the magnitude and direction of bias be predicted? This is a subject of ongoing research which is attempting to gather empirical evidence on factors (such as study design and intervention type) that determine the size and direction of these biases. The ability to predict both the likely magnitude of bias and the likely direction of bias would greatly improve the usefulness of evidence from systematic reviews of NRS. There is currently some evidence that in some limited circumstances the direction, at least, can be predicted (Henry 2001)

13.7.2 Evaluating the strength of evidence provided by reviews that include non-randomized studies

'Exposing' the evidence from NRS on a particular health question enables informed debate about its meaning and importance, and the certainty which can be attributed to it. Critically, there needs to be a debate about the chance that the observed findings could be misleading. Formal hierarchies of evidence all place NRS low down on the list, but above those of clinical opinion (Eccles 1996, National Health and Medical Research Council 1999, Oxford Centre for Evidence-based Medicine 2001). This emphasizes the general concern about biases in NRS, and the difficulties of attributing causality to the observed effects. The strength of evidence provided by a systematic review of NRS is likely to depend on meeting the challenges set out in Section 13.7.1. The ability to meet these challenges will vary with healthcare context and outcome. In some contexts little confounding is likely to occur. For example, little prognostic information may be known when infants are vaccinated, limiting possible confounding (Jefferson 2005).

Whether the debate concludes that there is a need for randomized trials or that the evidence from NRS is adequate for informed decision-making will depend on the cost placed on the uncertainty arising through use of potentially biased study designs, and the collective value of the observed effects. This value may depend on the wider healthcare context. It may not be possible to include assessments of the value within the review itself, and it may become evident only as part of the wider debate following publication.

For example, is evidence from NRS of a rare serious adverse effect adequate to decide that an intervention should not be used? The evidence is uncertain (due to a lack of randomized trials) but the value of knowing that there is the possibility of a potentially serious harm is considerable, and may be judged sufficient to withdraw the intervention. (It is worth noting that the judgement about withdrawing an intervention may depend on whether equivalent benefits can be obtained from elsewhere without such a risk; if not, the intervention may still be offered but with full disclosure of the potential harm.) Where evidence of benefit is not based on randomized trials and is therefore equivocal, the value attached to a systematic review of NRS of harm may be even greater.

In contrast, evidence of a small benefit of a novel intervention from a systematic review of NRS may not be sufficient for decision makers to recommend widespread implementation in the face of the uncertainty of the evidence and the substantial costs arising from provision of the intervention. In these circumstances, decision makers are likely to conclude that randomized trials should be undertaken if practicable and if the investment in the trial is likely to be repaid in the future.

The GRADE scheme for assessing the quality of a body of evidence is recommended for use in 'Summary of findings' tables in Cochrane reviews, and is summarized in Chapter 12 (Section 12.2). There are four quality levels: 'high', 'moderate', 'low' and 'very low'. A collection of studies that can be crudely categorized as randomized trials starts at the highest level, and may be downgraded due to study limitations (risk of bias), indirectness of evidence, heterogeneity, imprecision or publication bias. Collections of observational studies start at a level of 'low', and may be upgraded due to a large

magnitude of effect, lack of concern about confounders or a dose-response gradient. Review authors will need to make judgements about whether evidence from NRS should be upgraded from a low level or possibly (e.g. in the case of quasi-randomized trials) downgraded from a high level.

13.7.3 Guidance for potential review authors

Carrying out a systematic review of NRS is much more difficult than carrying out a systematic review of randomized trials. It is likely that complex decisions, requiring expert methodological or epidemiological advice, will need to be made at each stage of the review. Potential review authors should therefore seek to collaborate with epidemiologists or methodologists, irrespective of whether a review aims to investigate harms or benefits, short-term or long-term outcomes, frequent or rare events.

Healthcare professionals are keen to be involved in doing reviews of NRS in areas where there are few or no randomized trials because they have the ambition to improve the evidence-base in their specialty areas (the motivation for most Cochrane reviews). Methodologists are keen for more systematic reviews of NRS to inform the many areas of uncertainty in methodology highlighted by these chapters. However, healthcare professionals should also recognize that (a) the resources required to do a systematic review of NRS are likely to be much greater than for a systematic review of randomized trials and (b) the conclusions are likely to be much weaker and may make a relatively small contribution to the topic. Therefore, authors and CRG editors need to decide at an early stage whether the investment of resources is likely to be justified by the priority of the research question.

Bringing together the required team of healthcare professionals and methodologists may be easier for systematic reviews of NRS to estimate the effects of an intervention on long-term and rare adverse outcomes, for example when considering the side effects of drugs. However, these reviews may require the input of additional specialist authors, for example with relevant pharmacological expertise. There is a pressing need in many health conditions to supplement traditional systematic reviews of randomized trials of effectiveness with systematic reviews of adverse (unintended) effects. It is likely that these systematic reviews will usually need to include NRS.

13.8 Chapter information

Authors: Barnaby C Reeves, Jonathan J Deeks, Julian PT Higgins and George A Wells on behalf of the Cochrane Non-Randomised Studies Methods Group.

This chapter should be cited as: Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions.* Version 5.0.2 [updated September 2009]. The Cochrane Collaboration, 2009. Available from www.cochrane-handbook.org.

Acknowledgements: We gratefully acknowledge Ole Olsen, Peter Gøtzsche, Angela Harden, Mustafa Soomro, Guido Schwarzer and Bev Shea for their early drafts of different sections. We also thank Laurent Audigé, Duncan Saunders, Alex Sutton, Helen Thomas and Gro Jamtved for comments on previous drafts.

Box 13.8.a: The Cochrane Non-Randomised Studies Methods Group

The Non-Randomised Studies Methods Group (NRSMG) of the Cochrane Collaboration advises the Steering Group to set policy and formulate guidance about the inclusion of non-randomized studies (NRS) of the effectiveness of healthcare interventions in Cochrane reviews. Membership of the group is open to anyone who wishes to contribute actively to the work of group. The work of the group is primarily methodological, rather than focused on particular healthcare interventions.

Activities of NRSMG members include:

- Developing guidelines to help decide when to include non-randomized data in Cochrane reviews.
- Conducting methodological research in the use of non-randomized studies, including search methods, quality assessment, meta-analysis, pitfalls and misuse.
- Conducting empirical research to compare bias in systematic reviews using both randomized and non-randomized studies, and to identify conditions under which randomized and non-randomized studies have led to similar conclusions, and situations in which the conclusions have been clearly contradictory.
- Collating examples of healthcare questions that (a) have been studied using both non-randomized studies and randomized trials, and (b) have not been (or which for a long period have not been) studied adequately by means of randomized trials.
- Providing training at annual Cochrane Colloquia.

13.9 References

Audige 2004
Audige L, Bhandari M, Griffin D, Middleton P, Reeves BC. Systematic reviews of nonrandomized clinical studies in the orthopaedic literature. *Clinical Orthopaedics and Related Research* 2004: 249-257.

Chan 2004
Chan AW, Hróbjartsson A, Haahr MT, Gøtzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA* 2004; 291: 2457-2465.

Concato 1992
Concato J, Horwitz RI, Feinstein AR, Elmore JG, Schiff SF. Problems of comorbidity in mortality after prostatectomy. *JAMA* 1992; 267: 1077-1082.

Deeks 2003
Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, Petticrew M, Altman DG. Evaluating non-randomised intervention studies. *Health Technology Assessment* 2003; 7: 27.

Doll 1993
Doll R. Doing more good than harm: The evaluation of health care interventions: Summation of the conference. *Annals of the New York Academy of Sciences* 1993; 703: 310-313.

Downs 1998
Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology and Community Health* 1998; 52: 377-384.

Eccles 1996
Eccles M, Clapp Z, Grimshaw J, Adams PC, Higgins B, Purves I, Russel I. North of England evidence based guidelines development project: methods of guideline development. *BMJ* 1996; 312: 760-762.

Fraser 2006
Fraser C, Murray A, Burr J. Identifying observational studies of surgical interventions in MEDLINE and EMBASE. *BMC Medical Research Methodology* 2006; 6: 41.

Furlan 2006
Furlan AD, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant nonrandomized studies. *Journal of Clinical Epidemiology* 2006; 59: 1303-1311.

Glasziou 2007
Glasziou P, Chalmers I, Rawlins M, McCulloch P. When are randomised trials unnecessary? Picking signal from noise. *BMJ* 2007; 334: 349-351.

Golder 2006a
Golder S, Loke Y, McIntosh HM. Room for improvement? A survey of the methods used in systematic reviews of adverse effects. *BMC Medical Research Methodology* 2006; 6: 3.

Golder 2006b
Golder S, McIntosh HM, Duffy S, Glanville J, Centre for Reviews and Dissemination and UK Cochrane Centre Search Filters Design Group. Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. *Health Information and Libraries Journal* 2006; 23: 3-12.

Golder 2006c
Golder S, McIntosh HM, Loke Y. Identifying systematic reviews of the adverse effects of health care interventions. *BMC Medical Research Methodology* 2006; 6: 22.

Greenland 2004
Greenland S. Interval estimation by simulation as an alternative to and extension of confidence intervals. *International Journal of Epidemiology* 2004; 33: 1389-1397.

Grobbee 1997
Grobbee DE, Hoes AW. Confounding and indication for treatment in evaluation of drug treatment for hypertension. *BMJ* 1997; 315: 1151-1154.

Henry 2001
Henry D, Moxey A, O'Connell D. Agreement between randomized and non-randomized studies: the effects of bias and confounding. *9th Cochrane Colloquium*, Lyon (France), 2001.

Higgins 2003
Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; 327: 557-560.

Jefferson 2005
Jefferson T, Smith S, Demicheli V, Harnden A, Rivetti A, Di Pietrantonj C. Assessment of the efficacy and effectiveness of influenza vaccines in healthy children: systematic review. *The Lancet* 2005; 365: 773-780.

King 2005
King M, Nazareth I, Lampe F, Bower P, Chandler M, Morou M, Sibbald B, Lai R. Impact of participant and physician intervention preferences on randomized trials: a systematic review. *JAMA* 2005; 293: 1089-1099.

Kwan 2004
Kwan J, Sandercock P. In-hospital care pathways for stroke. *Cochrane Database of Systematic Reviews* 2004, Issue 2. Art No: CD002924.

MacLehose 2000
MacLehose RR, Reeves BC, Harvey IM, Sheldon TA, Russell IT, Black AM. A systematic review of comparisons of effect sizes derived from randomised and non-randomised studies. *Health Technology Assessment* 2000; 4: 1-154.

Moher 2001
Moher D, Schulz KF, Altman DG. The CONSORT Statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *The Lancet* 2001; 357: 1191-1194. (Available from www.consort-statement.org).

National Health and Medical Research Council 1999
National Health and Medical Research Council. *A guide to the development, implementation and evaluation of clinical practice guidelines [Endorsed 16 November 1998]*. Canberra (Australia): Commonwealth of Australia, 1999.

Oxford Centre for Evidence-based Medicine 2001
Oxford Centre for Evidence-based Medicine. Levels of Evidence [May 2001]. Available from: http://www.cebm.net/index.aspx?o=1047 (accessed 1 January 2008).

Peto 1995
Peto R, Collins R, Gray R. Large-scale randomized evidence: large, simple trials and overviews of trials. *Journal of Clinical Epidemiology* 1995; 48: 23-40.

Petticrew 2001
Petticrew M. Systematic reviews from astronomy to zoology: myths and misconceptions. *BMJ* 2001; 322: 98-101.

Pocock 2004
Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, Kasten LE, McCormack VA. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ* 2004; 329: 883.

Reeves 2004
Reeves BC, Gaus W. Guidelines for reporting non-randomised studies. *Forschende Komplementärmedizin und klassische Naturheilkunde* 2004; 11 Suppl 1: 46-52.

Reeves 2006
Reeves BC. Parachute approach to evidence based medicine: as obvious as ABC. *BMJ* 2006; 333: 807-808.

Reeves 2007
Reeves BC, Langham J, Lindsay KW, Molyneux AJ, Browne JP, Copley L, Shaw D, Gholkar A, Kirkpatrick PJ. Findings of the International Subarachnoid Aneurysm Trial and the National Study of Subarachnoid Haemorrhage in context. *British Journal of Neurosurgery* 2007; 21: 318-23.

Rothman 1986
Rothman KJ. *Modern Epidemiology*. Boston (MA): Little, Brown & Company, 1986.

Shadish 2002
Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston (MA): Houghton Mifflin, 2002.

Siegfried 2003
Siegfried N, Muller M, Volmink J, Deeks J, Egger M, Low N, Weiss H, Walker S, Williamson P. Male circumcision for prevention of heterosexual acquisition of HIV in men. *Cochrane Database of Systematic Reviews* 2003, Issue 3. Art No: CD003362.

Taggart 2001
Taggart DP, D'Amico R, Altman DG. Effect of arterial revascularisation on survival: a systematic review of studies comparing bilateral and single internal mammary arteries. *The Lancet* 2001; 358: 870-875.

Vandenbroucke 2007
Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, Egger M. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Medicine* 2007; 4: e297.

von Elm 2007
von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLoS Medicine* 2007; 4: e296.

Wells 2008
Wells GA, Shea B, O'Connell D, Peterson J, Welch V, Losos M, Tugwell P. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Available from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm (accessed 1 January 2008).

Wieland 2005
Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. *Journal of Clinical Epidemiology* 2005; 58: 560-567.