LITERATURE SEARCHING FOR RANDOMIZED CONTROLLED TRIALS USED IN COCHRANE REVIEWS: RAPID VERSUS EXHAUSTIVE SEARCHES

Pamela Royle
Ruairidh Milne
Wessex Institute for Health Research and Development

Abstract

Objectives: To analyze sources searched in Cochrane reviews, to determine the proportion of trials included in reviews that are indexed in major databases, and to compare the quality of these trials with those from other sources.

Methods: All new systematic reviews in the Cochrane Library, Issue1 2001, that were restricted to randomized controlled trials (RCTs) or quasi-RCTs were selected. The sources searched in the reviews were recorded, and the trials included were checked to see whether they were indexed in four major databases. Trials not indexed were checked to determine how they could be identified. The quality of trials found in major databases was compared with those found from other sources.

Results: The range in the number of databases searched per review ranged between one and twenty-seven. The proportion of the trials in the four databases were Cochrane Controlled Trials Register = 78.5%, MEDLINE = 68.8%, Embase = 65.0%, and Science/Social Sciences Citation Index = 60.7%. Searching another twenty-six databases after Cochrane Controlled Trials Register (CCTR), MEDLINE, and Embase only found 2.4% additional trials. There was no significant difference between trials found in the CCTR, MEDLINE, and Embase compared with other trials, with respect to adequate allocation concealment or sample size.

Conclusions: There was a large variation between reviews in the exhaustiveness of the literature searches. CCTR was the single best source of RCTs. Additional database searching retrieved only a small percentage of extra trials. Contacting authors and manufacturers to find unpublished trials appeared to be a more effective method of obtaining the additional better quality trials.

Keywords: Databases, Bibliographic; Technology assessment, Biomedical; Meta-analysis

Systematic reviews are increasingly used to help inform decisions about health care. Often the decision is about a treatment, when the most useful reviews are reviews of randomized controlled trials (RCTs) such as those prepared and disseminated by the Cochrane Collaboration (http://www.cochrane.org). It has been shown that Cochrane reviews have greater methodological rigor (11) and are better reported (10) than those published in paper-based journals. The validity of reviews is highly dependent on their including an unbiased sample

of relevant studies and so groups preparing Cochrane reviews pay particular attention to exhaustive search strategies (6). Such searches, however, are often time-consuming and costly; therefore, there has long been interest in the trade-off between timeliness and exhaustiveness when preparing reviews (22).

Questions about the value of exhaustive searches are particularly important for those preparing health technology assessment (HTA) reports for policy makers, who place special value on timeliness. We aimed to explore the trade-offs between timeliness and exhaustiveness by retrospectively analyzing literature searching in recent Cochrane reviews. Our specific objectives were to:

1. Determine the sources currently used in Cochrane reviews to search for trials.
2. Measure the proportion of trials indexed in four databases: CCTR (Cochrane Controlled Trials Register), MEDLINE, Embase, and S/SCI.
3. Analyze RCTs not indexed in CCTR, MEDLINE, or Embase.
4. Compare the quality of trials found in CCTR, MEDLINE, or Embase with those from other sources.

METHODS

Objective 1: To Determine the Sources Currently Used in Cochrane Reviews to Search for Trials

Systematic reviews were obtained from the Cochrane Database of Systematic Reviews (CDSR) contained in the Cochrane Library 2001 Issue 1. All reviews tagged "New" in that issue were selected, and the sections headed "Selection criteria" and the "Criteria for considering studies for this review: types of studies" were examined. Reviews were selected if they included only RCTs or quasi-RCTs (together referred to as "trials" in this study).

Information on the search strategy used for each review was taken from the section "Search strategy for identification of studies." Details of every source used was recorded in a spreadsheet. A source was defined as any of the following: a database; personal communication with experts or authors; a reference list; hand searching of journals or conference proceedings; contact with manufacturers. It was also noted whether the search was restricted by language.

Objective 2: To Measure the Proportion of Trials Indexed in Four Databases (CCTR, MEDLINE, Embase, and S/SCI)

All of the cited trials were checked to see whether they were indexed in each of the four major databases: Cochrane Controlled Trials Register (CCTR), MEDLINE, Embase, and Science Citation Index (SCI)/Social Sciences Citation Index (SSCI). (These latter two databases will be collectively referred to as S/SCI as they were searched together as a single database.) The databases were searched by using the first author's surname and a keyword in the title. Trials not found in the first search were looked for a second time by using different search terms, the name of a second author name and/or other title keywords.

Any trials that were not found in CCTR but were found in MEDLINE or Embase were checked to see whether they were RCTs or quasi-RCTs. The full articles of these trials were obtained and examined independently by both reviewers. The relevant sections of the foreign language articles were translated by the second author. Any differences in conclusions as to the study design were resolved by discussion.

Objective 3: To Analyze Trials not Indexed in CCTR, MEDLINE, Embase, or S/SCI

Trials that were not indexed in CCTR, MEDLINE, Embase, or S/SCI were checked to see whether they were indexed in the following twenty-five databases, chosen on the basis of

relevance and availability: AMED (Allied & Complementary Medicine), Article First, BIO-SIS Previews, BNI (British Nursing Index), CINAHL (Cumulative Index to Nursing & Allied Health Literature), Conference Papers Index, Contents First, Dissertation Abstracts, HMIC (Health Management Information Consortium), Index to Theses UK, Inside Information Plus, LILACS (Latin American and Caribbean Literature on the Health Sciences), Microbiology Abstracts, NRR (National Research Register, NTIS (National Technical Information Service), Papers First, PASCAL, Pedro (Physiotherapy Evidence Database), Proceedings First, PsychINFO, SIGLE (System for Information on Grey Literature in Europe), SportDISCUS, WoSP (Web of Science Proceedings), WorldCat, and Zetoc.

The following other details for trials indexed in any database were noted:

- Publication status. The four possible options were: (i) published data only, (ii) unpublished data only, (iii) a mixture of published plus unpublished data, or (iv) unpublished data were sought but not used;
- Publication format. For example, journal article, meeting abstract, conference proceeding, book chapter, thesis;
- Language.

Objective 4: To Compare the Quality of Trials Found in CCTR, MEDLINE, or Embase with Those in Other Sources

A quick proxy assessment of the quality of the RCTs was made by looking at two characteristics recorded in CDSR: the allocation concealment status of studies, and the number of patients they included.

(a) Checking the allocation concealment status of studies. A subgroup of reviews containing at least one trial not indexed in CCTR, MEDLINE, or Embase was studied. In cases where a study comprised more than one publication, and one was indexed in either CCTR, MEDLINE, or Embase (CME) and the other was not, the study was assigned to the category CME.

The allocation concealment assigned to each study was derived from the column labeled "Allocation concealment" in the table "Characteristics of Included Studies" included in every review. The four possibilities for allocation concealment status were adequate (A), unclear (B), inadequate (C), or that allocation concealment was not used as a criterion to assess validity (D). The relevant data for each trial were entered into a spreadsheet and then imported into Access and SPSS for statistical analysis.

(b) Comparison of patient numbers in CME versus non-CME trials. The subgroup of reviews that contained at least one trial not indexed in either CCTR, MEDLINE, or Embase were studied. The patient numbers for each trial were derived from the column labeled "Participants" in the table "Characteristics of Included Studies" included in every review.

RESULTS

Objective 1: The Sources Currently Used to Search in Cochrane Reviews for Trials

Figure 1 shows how the study sample of sixty-six reviews was derived. Nine of the reviews included no trials. The remaining 57 reviews included a total of 781 trials.

The sixty-six reviews searched a total of seventy-nine sources. The twenty most frequently searched are shown in Table 1. Nearly all reviews used some component of the Cochrane Collaboration's controlled trials register; but 3 of 66 (4.5%) of Cochrane
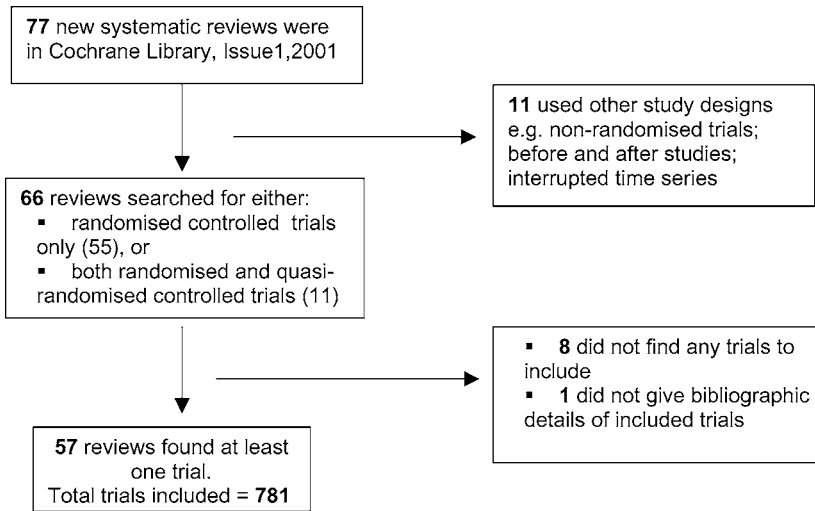
Figure 1. Derivation of the study sample.

reviews did not use any of these. MEDLINE (82%) and Embase (68%) were the next most commonly searched bibliographic databases. Eighty-two percent (54 of 66) of reviews sought either personal communication with authors or experts, or contact with a manufacturer.

The total number of sources used per review ranged between 2 and 30 (mean 8, median 6; interquartile range from 5 to 9). The number of databases searched per review ranged between 1 and 27 (mean, 5.5; median, 4; interquartile range from 3 to 6).

Table 1. Most Frequent Sources Searched in the 66 Reviews

| Rank	Source	Number of reviews in which searched
1	Medline	54 (81.8%)
2	Reference lists/bibliographies	53 (80.3%)
3	CCTR	49 (74.2%)
4	Embase	45 (68.2%)
5	Personal communication or contact with experts	44 (66.7%)
6	Collaborative Review Groups Trials Register	40 (60.6%)
7	Hand searching journals & conference papers	37 (56.1%)
8	Contacting manufacturers	25 (37.9%)
9	CINAHL	18 (27.3%)
10	SCI	16 (24.2%)
11	PsycLit	11 (16.7%)
12	SSCI	7 (10.6%)
13	AMED	6 (9.1%)
14	BIOSIS	6 (9.1%)
15	Dissertations Abstracts	6 (9.1%)
16	NRR	6 (9.1%)
17	HealthStar	5 (7.6%)
18	Web sites	5 (7.6%)
19	Cochrane Library	4 (6.1%)
20	Current Contents	4 (6.1%)
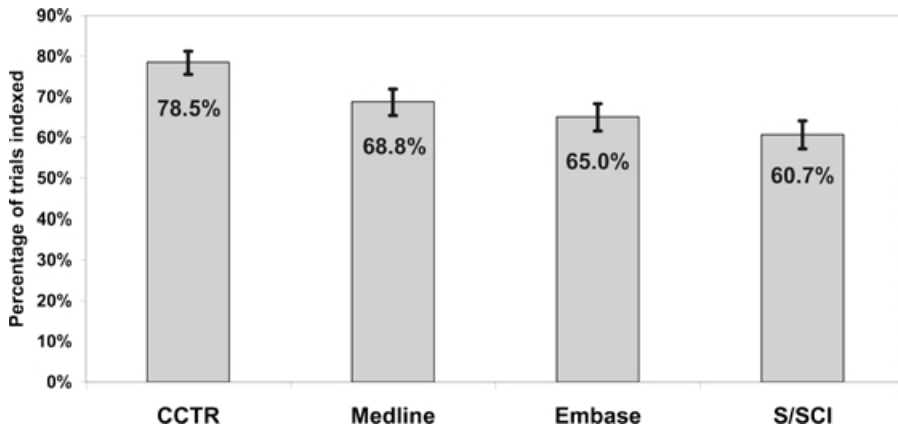
Figure 2. Percentages of trials indexed in the Cochrane Controlled Trials Register (CCTR), MEDLINE, Embase, and Science/Social Sciences Citation Index (S/SCI). Error bars represent 95% confidence intervals.

Language restrictions were acknowledged by two reviews (3%), whereas thirty (45%) stated there were none. Thirty-four reviews (52%) did not mention whether they had used language restrictions.

Objective 2: The Proportion of Trials That Were Indexed in CCTR, MEDLINE, Embase and S/SCI

We determined the proportion of the 781 trials that could be found in four databases: CCTR, MEDLINE, Embase, and S/SCI (Figure 2).

CCTR (78.5%) gave the highest coverage of trials for a single database, followed by MEDLINE (68.8%), Embase (65.0%) and S/SCI (60.7%). Note that the 95% confidence intervals for MEDLINE and Embase overlap, as do those for Embase and S/SCI, indicating that it is likely that a different sample of trials might yield a different rank order of these databases.

If reviewers were to search databases in ascending order of yield as measured in this study, the results would be as shown in Figure 3. The steps were as follows:

a) CCTR identified 78.5% of the 781 trials;

b) MEDLINE then yielded an extra twenty-six trials (3.3%);

c) Embase then identified an extra six trials (0.8%), four were non-English;

d) Science/Social Sciences Citation Index then yielded an extra seven trials (0.9%), and of these, three were journal articles (one non-English) and four were meeting abstracts.

Searching the further three databases after CCTR retrieved an extra thirty-nine (5%) additional trials. Note that the confidence intervals in Figure 3 are wide; repeating the process with a different sample of trials would not be expected always to produce the same rank order of databases.

We examined aspects of the overlap between the three CME databases.

a) The overlap between MEDLINE and Embase: The combined number of unique trials obtained from searching MEDLINE and then Embase was 583; of these, 462 were indexed in both databases.
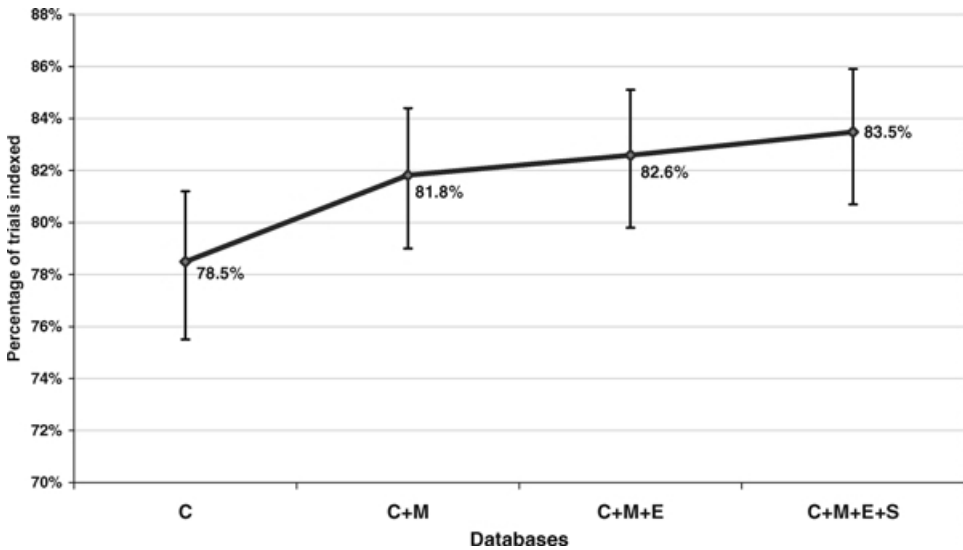
Figure 3. Cumulative percentage of trials found if databases are searched in the order Cochrane Controlled Trials Register (CCTR), MEDLINE, Embase, and Science/Social Sciences Citation Index (S/SCI). Error bars represent 95% confidence intervals. C, CCTR; M, MEDLINE; E, Embase; S, S/SCI.

Therefore, the overlap between these databases was 79% (95% confidence interval [CI], 76% to 82%); 36 of the 46 (78%) trials found in Embase but not in MEDLINE were non-drug trials.

b) Trials not in CCTR but in MEDLINE or Embase: Thirty-two "trials" not indexed in CCTR were found in either MEDLINE or Embase. Six of these (19%) proved to be neither RCTs nor quasi-RCTs, even though it was stated that these were the inclusion criteria for the review.

c) Trials in CCTR but not in MEDLINE or Embase: Sixty-two of the 781 trials (8%) were indexed in CCTR but not in MEDLINE and Embase. The publication types of these trials were 36 (59%) = conference proceedings or meeting abstracts, 12 (19%) = Chinese complementary medicine journals, 7 (11%) = book chapters, and 7 (11%) = other journal articles.

Objective 3: Trials not Indexed in CCTR, MEDLINE, or Embase

Trials not indexed in CCTR, MEDLINE, or Embase (non-CME trials) comprised 136 (17%) of the total of 781 trials and were found in 29 of the 57 reviews (51%). The distribution of the non-CME trials over these twenty-nine reviews was markedly skewed: thirteen of the twenty-nine reviews (45%) included only one non-CME trial, and 50% of the non-CME trials were derived from just two reviews, one based on an individual patient data (IPD) meta-analysis, *Tamoxifen for early breast cancer* (7), and one on Chinese medicinal herbs for chronic hepatitis B.

Of the 136 non-CME trials, 44 (32%) were categorized as "published," 16 (12%) as "published + unpublished," and 76 (56%) as "unpublished." Of the sixty non-CME trials categorized as either "published" or "published + unpublished," twenty-three (38%) were non-English (the majority being Chinese complementary medicine articles) and fourteen (23%) were meeting abstracts. There were insufficient details given for most of the "unpublished" trials to do a similar analysis.

Search strategies to optimize the yield of non-CME "published" or "published + unpublished" trials. Of the twenty-six databases searched for non-CME trials (S/SCI plus twenty-five others), only nineteen new trials were found; twelve of these were from BIOSIS and

Table 2. Allocation Concealments of the Trials in the 29 Reviews Containing at Least One Non-CME Trial

Allocation concealment	CME	Non-CME	Total
Adequate (A)	99 (29.6%)	26 (22.0%)	125
Unclear (B)	186 (55.7%)	35 (29.7%)	221
Inadequate (C)	41 (12.3%)	3 (2.5%)	44
Not used (D)	8 (2.4%)	54 (45.8%)	62
Total	334	118	452

Note: The difference between Adequate CME and non-CME trials, compared with the sum of all other categories (B + C + D) was calculated: Chi-squared = 2.522, d.f. = 1, p = .112. CME, trials indexed in either CCTR, MEDLINE, or Embase.

S/SCI. Of the additional twenty-four databases then searched, only seven additional trials were found. Those databases that were often searched in the reviews but found no trials not already indexed in CCTR, MEDLINE, or Embase were CINAHL, PsychLit, AMED, Pascal, and SIGLE.

The most effective search strategy would be one that took the least number of steps to identify the maximum number of unique trials, especially those outside the major databases. The optimum strategy to find the non-CME trials classified as either "published" or "published + unpublished" was first to check the bibliographies of other trials, and then to search BIOSIS and SCI, restricting the search to "meeting abstracts" or "book chapters" only.

Objective 4: Comparison of the Quality of Trials Found in CCTR, MEDLINE, and Embase with Those in Other Sources

a) Allocation concealment. The twenty-nine reviews which contained at least one non-CME trial were used as a source of trials to compare allocation concealment in CME and non-CME trials. The twenty-nine reviews contained 452 separate studies (some studies cited more than one trial). Table 2 shows that 30% of CME trials and 22% of non-CME trials were adequately concealed, but this difference was not significant (difference 7.6%, 95% CI 1.9% to 15.9%).

Nearly half (46%) of non-CME trials were classified as "D-allocation concealment not used as a criterion to assess validity" (Table 2). Nearly all of these (96%, 52 of 54) were from a single trial, an IPD meta-analysis (7). Table 3, therefore, makes a simpler comparison, looking at the frequency of "adequate" allocation concealment in those trials where this characteristic was used as a criterion to assess validity. It shows a lower proportion of "adequate" allocation concealment in the CME category compared with non-CME, but again this difference was not significant (difference, −10.3%; 95% CI, −23.4% to 2.1%).

Table 3. Proportion of Trials Classified with Allocation Concealments "Adequate" Compared with "Unclear or Inadequate"

Allocation concealment	CME	Non-CME	Total
Adequate (A)	99 (30.4%)	26 (40.6%)	125
Unclear or inadequate (B + C)	227 (69.6%)	38 (59.4%)	265
Total	326	64	390

Note: The difference between Adequate CME and non-CME trials compared with Unclear or Inadequate (B + C) was calculated: Chi-squared = 2.584, d.f. = 1, p = .108. CME, trials indexed in either CCTR, MEDLINE, or Embase.

Table 4. Patient Numbers in CME and non-CME Trials

	CME Trials ($n = 313$)	Non-CME trials ($n = 63$)
Mean	105.4	156.2
Median	60.0	74.0
SD	168.8	222.1
Range	9–1784	8–1309

Note: The *p* value of difference in patient numbers between two groups = .127 (from Mann Whitney *U* test). CME, trials indexed in either CCTR, MEDLINE, or Embase.

The publication formats of the twenty-six non-CME trials with adequate allocation concealment were eight (31%) = drug company reports, five (19%) = meeting abstracts, four (15%) = unpublished manuscripts in press or preparation, three (12%) = foreign language journal articles, two (8%) = ongoing trials, four (15%) = other formats. Fifteen of these twenty-six trials (58%) were classified as "unpublished."

b) Patient numbers. Table 4 shows patient numbers for the twenty-six reviews that contained at least one non-CME trial and for which patient numbers were available; this search gave a total of 376 studies. Although non-CME trials had a higher number of patients than CME trials, there was no significant difference between the two groups ($p = .127$). The publication status of the sixty-three non-CME trials was checked, and it was found that twenty-three (37%) were classified as unpublished, thirty (48%) as published, and ten (16%) as published + unpublished.

The characteristics of the larger non-CME trials, that is, those above median of seventy-four patients, were eight (25%) = meeting abstracts, eight (25%) = Chinese complementary medicine journal articles, five (16%) = manuscript in preparation or in press, five (16%) = drug company report, two (6%) = ongoing trial, one (3%) = new journal, one (3%) = Japanese journal article, one (3%) = German dissertation, one (3%) = book chapter.

DISCUSSION

This study of Cochrane reviews that were new in issue 1 of 2001 has shown that most searched between three and six databases in looking for trials, although the range (between one and twenty-seven databases) was wide. CCTR proved to be the single best source of trials (identifying 79% of the 781) with MEDLINE and Embase between them identifying an extra thirty-two trials (4.1%). There was a skewed distribution across the reviews for the 136 (17%) trials not found in CCTR, MEDLINE, or Embase, as 50% were found to come from just two reviews. There was no evidence that non-CME trials had poorer allocation concealment or smaller patient numbers than CME trials. The vast majority of the higher quality non-CME trials were either unpublished (drug company reports, meeting abstracts and unpublished manuscripts) or published in Chinese language complementary medicine journals.

The strengths of this study were that it analyzed a random sample of recent Cochrane reviews (so as to reflect current practice) and used rigorous methods, including independent duplicate checking of key "measurement" decisions. Also, a large number of alternative databases were thoroughly searched.

This study also had several limitations. It was a retrospective study, and it only analyzed a small number ($n = 66$) of reviews, thus increasing the probability a making a type II error; hence, this study would need to be tested in a larger sample of reviews to confirm these results. It also performed several statistical tests, thus increasing the chance that at least

one will be significant at the 5% level (and risk of type I error). Also, it was assumed that Cochrane reviewers fully and accurately reported their search strategies.

It was beyond the scope of this study to obtain the full articles of all the trials used in the reviews. Therefore, it had to rely on Cochrane reviewers' classifications as to the publication status and allocation concealment of trials, and it is probable that there were some inconsistencies between different reviewers. In addition, it is not known how accurate allocation concealment and patient numbers are as proxy measures of trial quality. Also, it measured the outcome of searching in terms of numbers of unique trials found and not in terms of the information they added (the clearest example of which would be a change in a meta-analyzed result). Finally, this study did not look at costs, in terms of information scientist time and reviewer time.

Previous research relevant to this study has compared the comprehensiveness of MEDLINE with hand searching (1;17) or with searches of other databases, including Embase, CINAHL, BIOSIS, PsychLit, and LILACS (2;3;4;9;16;19;24). One study looked at unpublished dissertations as a source of trials for systematic reviews (25). A recent study has compared trials indexed in MEDLINE with trials not indexed in MEDLINE, and found that the latter overestimated treatment effects in meta-analyses by approximately 5% (8). Also, research has investigated the impact of grey literature and non-English trials on the results of meta-analyses (13;18).

However, we are not aware of prior studies similar to this one that have compared the comprehensiveness of searches of CCTR with other databases. Also, none have quantified what is lost in terms of the number and quality of trials found when limiting searching for Cochrane reviews to a few databases and, hence, have allowed some estimate of the extent to which CCTR has made it possible to do searches for RCTs that are both rapid and comprehensive.

This study provided some insight into the "added value" of searching CCTR over MEDLINE and Embase. Eight percent (62 of 781) of trials in CCTR were not found in MEDLINE or Embase. The majority of these reports were conference proceedings or meeting abstracts (59%) and articles from hand searching Chinese complementary medicine journals (19%). It may also be that the sixty-two extra trials found in CCTR and not in MEDLINE or Embase were only there because the reviewers had identified them when searching for the review and forwarded them to Cochrane for inclusion in CCTR. However, this method is very unlikely to account for most of the trials, as 66% (41 of 62) were from journals hand searched by the Cochrane Collaboration and 68% (42 of 62) were in CCTR 2000 issue1 (a year before the review being published).

It may be that the true percentage of "CCTR-only" trials (i.e., those in CCTR but not in MEDLINE or Embase) is actually higher than 8%, but they were under-represented in these reviews. Possible reasons for this discrepancy could be that (i) the "CCTR-only" trials were not judged of sufficient quality to be included in the reviews, (ii) many are non-English and were excluded due to difficulties and expense of obtaining translations, and (iii) they were not relevant to the subjects being reviewed here. Conversely, the true percentage of "CCTR-only" trials in CCTR could be lower than 8%, and they were over-represented in this study as they were particularly relevant to the subjects reviewed here.

If we exclude studies that proved not to be trials, searching MEDLINE and Embase after CCTR had been searched only yielded an extra 3.7% (20 of 537) trials from MEDLINE, and 1.2% (6 of 508) from Embase. As MEDLINE and Embase are the primary source for trials for CCTR, this is not surprising. After searching CCTR, it appears that there is very little extra yield for the large amount of time invested in searching MEDLINE and Embase. The only exception to this would be a search for the most recent six to twelve months, for those trials that have not had time to be included in CCTR.

The overlap between MEDLINE and Embase has been estimated variously at between 10% and 75%, depending on the topic being searched (6;23). In this study, the overlap between MEDLINE and Embase for trials was found to be 79.2%. This higher percentage may be because this study only measured the overlap between trials that have been included in reviews, whereas the others either included a broader range of studies or were not restricted to only trials that had been included in reviews.

In 49% (28 of 57) of the reviews, all of the included trials could be found by searching only three databases: CCTR, MEDLINE and Embase. Hence, there were twenty-nine reviews that contained at least one trial not in any of these databases. The distribution of non CME-trials was very unevenly distributed across the twenty-nine reviews, with thirteen reviews each containing only one trial not in CME and just two reviews accounting for 50% of the non-CME trials.

Therefore, in most of the reviews, not searching beyond the three databases would have made little difference to the number of trials included. However, it is not known whether excluding the non-CME trials would have made a difference to the results of the review, as we do not know if the size and direction of treatment effects are systematically different in CME versus non-CME trials.

The law of diminishing returns became quickly apparent when searching beyond BIOSIS and SCI for new trials, as searching twenty-four extra databases only retrieved an additional seven trials. None of the following databases contributed any additional trials: CINAHL, PsychLit, AMED, PASCAL, and SIGLE.

Over half (56%) of the trials not in CCTR, MEDLINE, or Embase were classified as unpublished, so they would not be found by searching databases. Checking the bibliographies of the more accessible trials appeared to be an effective method of identifying additional relevant trials. However, it is acknowledged that checking of citations in articles does have the danger of introducing citation bias into reviews.

It was of some concern that six of thirty-two (19%) of the trials examined were included in reviews of RCTs, when an examination of the full article revealed that they clearly were not RCTs. These six trials were spread over four reviews, indicating that the problem was not due to only one reviewer misclassifying trials.

The Cochrane Reviewers Handbook (5) states the quality of the trials used in reviews should be assessed, although there is no widespread consensus on how this should be done; however, it is generally thought that concealment of treatment allocation, blinding of outcome assessment, and handling of patient attrition in the analysis should be assessed (12;15). Results from a meta-analysis (8) have shown that trials with inadequate or unclear allocation concealment are associated with an exaggeration of treatment effects of approximately 30%, and inadequate allocation concealment appears to be a more significant source of bias than the other aspects of trial quality investigated so far.

Results from this study indicated that most of the non-CME trials with adequate allocation concealment were unlikely to be found in databases, as they were either (i) unpublished (e.g., drug company reports, manuscripts in press or in preparation or ongoing trials, or (ii) indexed to only a very limited extent (e.g., meeting abstracts or foreign language journals). Therefore, to locate them, one would need to contact drug companies and hand search meeting abstracts and relevant foreign language journals.

Smaller studies are thought to be associated with lower methodological quality and to be more susceptible to publication bias than larger trials (14;15;20;21). The average number of patients in trials not indexed in CCTR, MEDLINE, and Embase was higher than those that were indexed in these databases (although this difference was not statistically significant). As larger trials have a bigger impact on overall estimates of treatment effects in a meta-analysis than smaller trials, it is more important that these larger trials are not missed.

Box 1. Recommended Search Strategy for Optimizing the Retrieval of RCTs from the Literature

Step	Comment
1. Search CCTR	This should retrieve about 79% of the total trials, which would be found if an exhaustive search were to be done
2. Search the records added to MEDLINE (via PUBMED) and Embase in the previous year	This is because there is an inevitable time lag before new trials added to these databases appear in CCTR
3. Search BIOSIS and SCI, restricted to meeting abstracts only	This is because both these databases, unlike MEDLINE or Embase, index the individual meeting abstracts that appear in the journal supplements indexed by these databases
4. Check the bibliographies of the relevant trials found so far	However, the possibility of citation bias should be borne in mind
5. Contact authors of major trials and abstracts identified, and experts in the area, for unpublished data	More than half (56%) of the trials not found in the major databases in this study were unpublished, and on the basis of the data presented here, there is no prima facie reason to exclude them on the basis that they are of poorer methodological quality or smaller sample size; also, it is known that about 50% of trials published in abstract form are never published in full
6. Contact manufacturers if the review is of a drug or device	
7. Search other databases, including grey literature databases	But expect that this may result in very few new trials being identified
8. Hand search key journals	Particularly if the review is of complementary medicine
9. Do a sensitive search of Medline and Embase for RCTs missed by the Cochrane Collaboration and not in CCTR	

RCT, randomized controlled trial.

The publication formats of the largest non-CME trials, that is, those above the median number of seventy-four, showed that they were most commonly meeting abstracts (25%), Chinese articles on complementary medicine (25%), or drug company reports (16%). As such trials are unlikely to be included in databases, this finding suggests that hand searching meeting abstracts and contacting experts and manufacturers of drugs is important to ensure a comprehensive search. Hand searching of journals is probably only productive in the case of complementary medicine reviews.

Those commissioning HTA reports and those searching for trials to contribute to such reports need to be aware of the law of diminishing returns demonstrated in this study. Based on the results of this study, we would suggest the search strategy shown in the Box 1, although the small size of this study means that it can only be a suggestion. We tentatively make the following observations to information scientists working within the Cochrane Collaboration (Box 2).

Box 2. Observations to the Cochrane Collaboration

• Abstracts and conference proceedings could be a productive source of trials for indexing in CCTR
• New Cochrane reviews could be a useful source of new RCTs for hand searchers
• There are inconsistencies in the reporting of language restrictions in Cochrane Reviews
• There are very wide variations between Cochrane reviews in their searching: whether they search CCTR and how many further sources they search; this produces inconsistencies as well as some limited searching

RCTs, randomized controlled trials.

This study has looked only at searching for RCTs for Cochrane reviews and does not address searching for other sort of reviews and other study designs. Therefore, future research should analyze systematic reviews published outside the Cochrane Library and also reviews that include other study designs, such as nonrandomized trials, cohort studies, and case series. Also, it would be interesting to determine how much difference the exclusion of trials not indexed in major databases have on the final results of a systematic review. Although one such study has been done (8), it would be useful to repeat this with other databases and over a wider range of subject areas, including complementary medicine.

REFERENCES

1. Adams CE, Power A, Frederick K, Lefebvre C. An investigation of the adequacy of MEDLINE searches for randomized controlled trials (RCTs) of the effects of mental health care. *Psychol Med*. 1994;24:741-748.
2. Brazier H, Begley CM. Selecting a database for literature searches in nursing: MEDLINE or CINAHL? *J Adv Nurs*. 1996;24:868-875.
3. Burnham J, Shearer B. Comparison of CINAHL, EMBASE, and MEDLINE databases for the nurse researcher. *Med Ref Serv Q*. 1993;12:45-57.
4. Castro AA, Clark OA, Atallah AN. Optimal search strategy for clinical trials in the Latin American and Caribbean Health Science Literature Database (LILACS). *Rev Paul Med*. 1997;115:1423-1426.
5. Clarke M, Oxman AD, eds. Assessment of study quality. Cochrane Reviewers Handbook 4.1.2 [updated March 2001] Section 6. In: *The Cochrane Library*. Oxford: Update Software, 2001 Issue 2.
6. Clarke M, Oxman AD, eds. Locating and selecting studies. Cochrane Reviewers Handbook 4.1.2 [updated March 2001] Section 5. In: *The Cochrane Library*. Oxford: Update Software, 2001 Issue 2.
7. Early Breast Cancer Trialists' Collaborative Group. Tamoxifen for early breast cancer (Cochrane Review). In: *The Cochrane Library*. Oxford: Update Software, 2001 Issue 1.
8. Egger M, Ebrahim S, Smith GD. Where now for meta-analysis? *Int J Epidemiol*. 2002;31:1-5.
9. Fergusson D, Laupacis A, Salmi LR, et al. What should be included in meta-analyses? An exploration of methodological issues using the ISPOT meta-analyses. *Int J Technol Assess Health Care*. 2000;16:1109-1119.
10. Jadad AJ, Moher D, Browman GP, et al. Systematic reviews and meta-analysis on treatment of asthma: Critical evaluation. *BMJ*. 2000;320:537-540.
11. Jadad AR, Cook DJ, Jones A, et al. Methodology and reports of systematic reviews and meta-analyses: A comparison of Cochrane reviews with articles published in paper-based journals. *JAMA*. 1998;280:278-280.
12. Juni P, Altman DG, Egger M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42-46.
13. Juni P, Holenstein F, Sterne J, et al. Direction and impact of language bias in meta-analyses of controlled trials: Empirical study. *Int J Epidemiol*. 2002;31:115-123.
14. Kjaergard LL, Nikolova D, Gluud C. Randomized clinical trials in hepatology: Predictors of quality. *Hepatology*. 1999;30:1134-1138.
15. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in meta-analyses. *Ann Intern Med*. 2001;135:982-989.
16. Kleijnen J, Knipschild P. The comprehensiveness of MEDLINE and EMBASE computer searches. Searches for controlled trials of homoeopathy, ascorbic acid for common cold and ginkgo biloba for cerebral insufficiency and intermittent claudication. *Pharm Weekbl Sci*. 1992;14:316-320.
17. Langham J, Thompson E, Rowan K. Identification of randomized controlled trials from the emergency medicine literature: Comparison of hand searching versus MEDLINE searching. *Ann Emerg Med*. 1999;34:25-34.
18. McAuley L, Pham B, Tugwell P, Moher D. Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analyses? *Lancet*. 2000;356:1228-1231.

19. McDonald S, Taylor L, Adams C. Searching the right database. A comparison of four databases for psychiatry journals. *Health Libr Rev*. 1999;16:151-156.

20. Sterne JA, Egger M, Davey SG. Investigating and dealing with publication and other biases. In: Egger M, Davey SG, Altman DG, eds. *Systematic reviews in health care: Meta-analysis in context*. London: BMJ Books; 2001:189-208.

21. Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: Power of statistical tests and prevalence in the literature. *J Clin Epidemiol*. 2000;53:1119-1129.

22. Stevens A, Colin-Jones A, Gabbay J. "Quick and clean": Authoritative health technology assessment for local health care contracting. *Health Trends*. 1995;27:37-42.

23. Suarez-Almazor ME, Belseck E, Homik J, et al. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. *Control Clin Trials*. 2000;21:476-487.

24. Topfer LA, Parada A, Menon D, et al. Comparison of literature searches on quality and costs for health technology assessment using the MEDLINE and EMBASE databases. *Int J Technol Assess Health Care*. 1999;15:297-303.

25. Vickers AJ, Smith C. Incorporating data from dissertations in systematic reviews. *Int J Technol Assess Health Care*. 2000;16:711-713.