A comparison of handsearching versus MEDLINE searching to identify reports of randomized controlled trials

S. Hopewell*,†, M. Clarke, A. Lusher, C. Lefebvre and M. Westby

The U.K. Cochrane Centre, Oxford, U.K.

SUMMARY

This study aims to compare handsearching to a basic MEDLINE search for the identification of reports of randomized trials in specialized health care journals. Twenty-two specialized health care journals, published in the U.K., were handsearched for all reports of controlled trials (as defined by the Cochrane Collaboration). The reports of trials, which were judged to be definitely randomized, were identified from a random sample of three years per journal and form one element of this study. A MEDLINE search using the publication type terms ‘randomized controlled trial’ and ‘controlled clinical trial’ was also performed for the same journal years. The reports of trials retrieved by handsearching were then compared against those retrieved from the MEDLINE search, to identify differences in retrieval between the two techniques. Reports of randomized trials identified by the MEDLINE search but not found by handsearching were individually assessed to see if they met the Cochrane eligibility criteria for a report of a randomized trial. A total of 714 reports of randomized trials were found by using a combination of both handsearching and MEDLINE searching. Of these, 369 (52 per cent) were identified only by handsearching and 32 (4 per cent) were identified only by MEDLINE searching. Of the reports identified only by handsearching, 252 had no MEDLINE record, of which 232 (92 per cent) were meeting abstracts or published in supplements; 117 (25 per cent) of the 462 reports of randomized trials which had a MEDLINE record were missed by the electronic search because they did not have either of the publication type terms ‘randomized controlled trial’ or ‘controlled clinical trial’. This proportion varied depending on when the reports of randomized trials were published (that is, before or after the introduction of the MEDLINE publication type terms above). The highest additional yield from handsearching compared to MEDLINE searching was for reports of randomized trials published prior to 1991 and from handsearching the non-MEDLINE indexed parts of a journal. The results of this study suggest that a combination of MEDLINE and handsearching is required to identify adequately reports of randomized trials. © Crown copyright 2002. Reproduced with the permission of Her Majesty’s Stationery Office. Published by John Wiley & Sons, Ltd.

KEY WORDS: Cochrane Collaboration; MEDLINE; randomized controlled trials; information retrieval; handsearching

*Correspondence to: Sally Hopewell, The U.K. Cochrane Centre, Summertown Pavilion, Middle Way, Oxford OX2 7LG, U.K.
†E-mail: shopewell@cochrane.co.uk

Contract/grant sponsor: European Union BIOMED Programme; contract/grant number: BMH4-CT98-3803

© Crown copyright 2002. Reproduced with the permission of Her Majesty’s Stationery Office. Published by John Wiley & Sons, Ltd.
INTRODUCTION

Randomized controlled trials, involving sufficient numbers of participants, are essential to distinguish reliably between the effects of health care interventions and the effects of bias or chance. The dissemination of the trials’ results through systematic reviews of the findings is just as important as conducting the trials and provides a basis for rational decision-making about the effects of health care interventions [1]. The Cochrane Collaboration is an international organization which is committed to preparing, maintaining and promoting the accessibility of systematic reviews of health care interventions [2].

The validity of the results of a systematic review is highly dependent on the results of the underlying data and it is, therefore, necessary to identify as unbiased and complete a set of relevant studies as possible. Trial identification can be a particularly time-consuming part of the systematic review process and the Cochrane Collaboration has done a large amount of work to identify relevant studies [3]. In addition, in the absence of a prepared systematic review, individuals might rely on a quick and simple search of MEDLINE to identify reports of randomized trials. Even though many reports of studies are included in bibliographic databases, such as MEDLINE and EMBASE, the majority of journals are not indexed in these databases, especially journals published in languages other than English. For example, a study of 68 Spanish journals in general medicine found that only six were indexed in MEDLINE [4].

Even when journals are indexed in databases such as MEDLINE, it can still be difficult to identify all relevant studies. This can be for a number of reasons, one of which is the lack of availability of appropriate terms to the U.S. National Library of Medicine indexers. The MEDLINE publication type term for ‘randomized controlled trial’ was only introduced in 1991; likewise the publication type term ‘controlled clinical trial’ was not introduced until 1995. Some research has also suggested an inconsistency in indexing by MEDLINE indexers [5]. Lack of cover-to-cover indexing is also a potential problem, as some sections of a journal might not be included in MEDLINE. This may be a particular problem with supplements and conference abstracts [6]. Finally, authors may not have described their research methods clearly enough to allow accurate indexing of the methodology.

To address these problems, over 1700 journals have been or are being handsearched within the Cochrane Collaboration to identify reports of controlled trials (www.cochrane.org/cochrane/hsearch.htm). The reports of identified trials form part of an international register of controlled trials published as The Cochrane Controlled Trials Register (central), which is available in The Cochrane Library [7]. Reports found by handsearching that are inappropriately indexed in MEDLINE may be sent to the U.S. National Library of Medicine for re-tagging with appropriate indexing terms.

In addition, a highly sensitive search strategy for identifying reports of controlled trials in MEDLINE was developed in 1993 by one of the authors (CL) [8]. The first and second phases of this search strategy have been run against MEDLINE, the titles and abstracts have been assessed and those records judged to be trial reports where participants were definitely or possibly randomized to receive health care interventions have been submitted to the U.S. National Library of Medicine for re-tagging with appropriate indexing terms. At the time of this study, this had been carried out for the period 1966 to 1997 [3].
AIMS

The aim of this study was to compare handsearching to a basic MEDLINE search for the identification of reports of randomized trials in specialized health care journals.

This study draws on data derived from an international project, co-ordinated by two of the authors (CL and AL), undertaken by European Cochrane Centres to handsearch specialized health care journals published in Western Europe for reports of controlled trials.

METHODS

For this study 22 specialized health care journals searched by the U.K. Cochrane Centre handsearchers were selected. The inclusion criteria for journals were that they were known to contain a high proportion of reports of controlled trials, known to be indexed in MEDLINE, were published in the U.K. and were accessible locally for handsearching.

Handsearching

The 22 journals were handsearched for all reports of controlled trials, as defined by the Cochrane Collaboration. This definition requires that ‘the participants (or other units) were definitely or possibly assigned prospectively to one of two (or more) alternative forms of health care using a process of random allocation’. To search a journal by hand, a trained person is required to check a journal from cover to cover, reading each item until he or she is satisfied as to whether or not it is an eligible report, according to the criteria above, irrespective of the health care intervention or condition under investigation. All full reports, short reports, editorials, correspondence sections, meeting abstracts and supplements are checked. All handsearchers in this study had received training on identifying reports of controlled trials by U.K. Cochrane Centre staff. All reports that were judged to be potentially relevant were then individually assessed by experienced handsearchers (MC or MW). This was to confirm that they met the Cochrane eligibility criteria and to act as a quality control measure to minimize over-inclusion. Random checks of the journals handsearched were also made (a sample of approximately 10 per cent) as a further quality control measure to minimize under-inclusion. All reports of trials identified by the handsearching have subsequently been submitted for inclusion in The Cochrane Controlled Trials Register (central) and, where appropriate, submitted for re-tagging in MEDLINE with the relevant publication type term.

For the purposes of this study, only those trial reports where there was sufficient information to conclude that the participants were randomized to different health care interventions were analysed. A sample of three years per journal were selected at random from within the sample of years available at the time of conducting this study, giving a total of 66 journal years, ranging from 1970 to 1998. A list of journals and details of the years selected for this study are given in Tables I and II.

MEDLINE searching

A MEDLINE search using PubMed was carried out for the same journal years in October 1999 (www.ncbi.nlm.nih.gov/entrez/query.fcgi). The search used the publication type terms...
Table I. List of journals handsearched.

<table>
<thead>
<tr>
<th>Journal title</th>
<th>Years assessed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bone Marrow Transplantation</td>
<td>1986, 1994, 1998</td>
</tr>
<tr>
<td>European Journal of Surgical Oncology</td>
<td>1975, 1985, 1995</td>
</tr>
<tr>
<td>Pharmatherapeutica</td>
<td>1976, 1984, 1988</td>
</tr>
</tbody>
</table>

Table II. Number of journal years in each time period.

<table>
<thead>
<tr>
<th>Years assessed</th>
<th>Number of journal years</th>
</tr>
</thead>
<tbody>
<tr>
<td>1970–1990</td>
<td>35</td>
</tr>
<tr>
<td>1991–1997</td>
<td>24</td>
</tr>
<tr>
<td>1998</td>
<td>7</td>
</tr>
</tbody>
</table>

‘randomized controlled trial’ and ‘controlled clinical trial’ (which are referred to in Figure 3 as RCT/CCT in PT) combined with the journal title and year of publication.

A copy of the original paper was obtained for reports identified by the MEDLINE search but not found by the handsearching. These were individually assessed by one of the authors (MW, who is an experienced handsearcher) to see if they met the Cochrane eligibility criteria. For the purposes of this study, only those trial reports where there was sufficient information to conclude that the participants were randomized to different health care interventions were included in this analysis.

Comparison of handsearching and MEDLINE searching

Reports of randomized trials identified by either method (handsearching or MEDLINE searching) were compared and differences in retrieval between the two methods were assessed. This
report summarizes the overall findings and presents data for the reports published during three
time periods:

(i) 1970–1990, when the assignment of the publication type term ‘randomized controlled
trial’, to these particular records, would be due to re-tagging in MEDLINE;
(ii) 1991–1997, when the assignment of the publication type term ‘randomized controlled
trial’, to these particular records, could be due to either re-tagging in MEDLINE or
prospective MEDLINE indexing;
(iii) 1998, when the assignment of the publication type term ‘randomized controlled trial’,
to these particular records, would be due to prospective MEDLINE indexing.

RESULTS

A total of 66 journal years were assessed. The findings from both methods (handsearching
and MEDLINE searching) were combined and 714 reports of randomized trials were found
that were eligible for inclusion in The Cochrane Controlled Trials Register (central). Of these, 369/714 (52 per cent) were identified only by handsearching and had been missed by
the MEDLINE search and 32/714 (4 per cent) were identified only by the MEDLINE search
and had been missed by handsearching. A total of 21 of the 366 (6 per cent) records indexed
with the publication type term ‘randomized controlled trial’ in MEDLINE were discarded
following assessment, as they were not judged to be reports of randomized trials.

Of the 369 reports of randomized trials identified only by handsearching, 252/369 (68
per cent) had no MEDLINE record (232/252 (92 per cent) of these were abstracts and/or
published in supplements).

A total of 462 reports of randomized trials had a MEDLINE record, of which 117/462 (25
per cent) were not indexed with the publication type term ‘randomized controlled trial’ or
‘controlled clinical trial’. The remaining 345 reports did have a MEDLINE record indexed
with one of these terms.

Differences over time

Reports of randomized trials published during the three specified time periods were analysed
separately (Figure 1).

1970 to 1990

For the period 1970 to 1990, 35 journal years were assessed. By combining the results of
both methods (handsearching and MEDLINE), 212 reports of randomized trials were found.
Of these, 104/212 (49 per cent) were identified only by handsearching and had been missed by
the MEDLINE search and 3/212 (1 per cent) were identified only by the MEDLINE search
and had been missed by handsearching.

Of the 104 reports of randomized trials identified only by handsearching, 26/104 (25 per
cent) had no MEDLINE record (13/26 (50 per cent) were abstracts and/or published in
supplements).

A total of 186 reports of randomized trials had a MEDLINE record, of which 78/186 (42
per cent) were not indexed with the publication type term ‘randomized controlled trial’ or
For the period 1991 to 1997, 24 journal years were assessed. By combining the results of both methods, 400 reports of randomized trials were found. Of these, 211/400 (53 per cent) were identified only by handsearching and had been missed by the MEDLINE search and 17/400 (4 per cent) were identified only by the MEDLINE search and had been missed by handsearching.

Of the 211 reports of randomized trials identified only by handsearching, 180/211 (85 per cent) had no MEDLINE record (176/180 (98 per cent) were abstracts and/or published in supplements).

A total of 220 reports of randomized trials had a MEDLINE record, of which 31/220 (14 per cent) were not indexed with the publication type term ‘randomized controlled trial’ or ‘controlled clinical trial’ (the remaining 189 had a MEDLINE record indexed with one of these terms).

Seven journal years were assessed for the period 1998. By combining the results of both methods, 102 reports of randomized trials were found. Of these, 54/102 (53 per cent) were identified only by handsearching and had been missed by the MEDLINE search and 12/102 (12 per cent) were identified only by the MEDLINE search and had been missed by handsearching.

Of the 54 reports of randomized trials identified only by handsearching, 46/54 (85 per cent) had no MEDLINE record (43/46 (93 per cent) were abstracts, all of which were published in supplements).
A total of 56 reports of randomized trials had a MEDLINE record, of which 8/56 (14 per cent) were not indexed with the publication type term ‘randomized controlled trial’ or ‘controlled clinical trial’ (the remaining 48 had a MEDLINE record indexed with one of these terms).

DISCUSSION

One of the reasons for the study reported here was to assess the adequacy of a simple two-term search strategy to identify reports of randomized trials relative to handsearching. Preliminary searches have identified at least 50 other studies comparing handsearching to MEDLINE searching. The results of those studies will be reported in more detail as part of a Cochrane Methodology Review [9]. However, most of these studies have used more complex and highly sensitive search strategies than the methods that have been used here.

Our results show that handsearching more than doubled the number of reports of randomized trials found using the basic MEDLINE search. Of particular interest are the reasons why the MEDLINE search failed to identify just over 50 per cent of the reports of randomized trials found by handsearching.

No MEDLINE record

The results of this study suggest that one of the main reasons for failure to identify reports of randomized trials in MEDLINE is simply the absence of the relevant record in MEDLINE. More than half (68 per cent) of the reports of randomized trials found only by handsearching had no MEDLINE record. However, 92 per cent of these were conference abstracts and/or published in journal supplements, which are not routinely indexed in MEDLINE (Figure 2). Obviously some of the conference abstracts identified will subsequently have been published as full MEDLINE-indexed papers. However, there is strong evidence to suggest that only approximately half of all conference abstracts describing the results of trials are subsequently ever published in full [10].

These results are consistent with earlier studies, which suggest that a major reason for failure to identify reports of randomized trials in MEDLINE is simply because they are published as conference abstracts or in supplements that are not routinely indexed in MEDLINE [6, 11–13]. Currently the U.S. National Library of Medicine indexes all articles, substantive editorials and letters. However, they do not index conference abstracts, book reviews, software or equipment reviews, meeting announcements, or other non-article items from a journal [14]. They do, however, index some conference abstracts which are published in special supplements. This has important implications as research suggests that trials reported in non-indexed parts of journals are likely to have systematically different results from trials published as full, MEDLINE-indexed, papers [15].

Lack of appropriate MEDLINE indexing terms

The other important reason for failure to identify reports of randomized trials in MEDLINE was a lack of the relevant MEDLINE indexing terms (Figure 3). For the period before 1991, when neither the publication type term ‘randomized controlled trial’ nor ‘controlled clinical trial’ had been available to MEDLINE indexers, 42 per cent of reports of randomized trials
that had a MEDLINE record were not indexed with either of these terms. This figure dropped considerably (to only 14 per cent) in recent years, suggesting that the ability to identify reports of randomized trials in MEDLINE, even using a simple search strategy, has improved.
considerably since the introduction of the publication type terms ‘randomized controlled trial’ in 1991 and ‘controlled clinical trial’ in 1995. These results support those of an earlier study, which used the same simple two-term search strategy to identify reports of randomized trials in the obstetrics and gynaecology literature [16]. The findings of that study also showed that the sensitivity and precision of this basic MEDLINE search had improved over time, especially in recent years.

Our results also suggest that the work of the Cochrane Collaboration and the U.S. National Library of Medicine in the re-tagging of MEDLINE has had an important impact on the ease of identifying reports of randomized trials. For the period before 1991, 58 per cent of the reports of randomized trials that had a MEDLINE record had been given the publication type term ‘randomized controlled trial’ or ‘controlled clinical trial’, as a result of this collaborative effort to re-tag MEDLINE.

**Inadequacy of handsearching**

A small number (4 per cent) of reports of randomized trials were missed by handsearching. If, however, a more complex, highly sensitive MEDLINE search had been used in our study, we might have identified additional reports of randomized trials that had been missed by the handsearching. For example, an earlier study using a complex MEDLINE search found that handsearchers missed 17 per cent of reports of randomized trials [6].

Finally, one should consider the effect that the choice of journal years assessed (that is, three years per journal) might have had on the results of this study. It is possible that the choice of years assessed could have introduced an element of bias into this study as they may not have been representative of the whole journal. It is hoped, however, that this has been minimized by the random selection of the years to be assessed.

**IMPLICATIONS FOR PRACTICE AND RESEARCH**

Handsearching still has a valuable role to play in attempting to identify an unbiased or complete set of reports of randomized trials for inclusion in systematic reviews of health care interventions. However, for these U.K. specialized journals at least, the findings of this study suggest that, in addition to a basic MEDLINE search, the highest additional yield of reports of randomized trials is from handsearching the non-MEDLINE indexed parts of a journal and journals published pre-1991. Where resources are limited, the results of this study would suggest that handsearching is best targeted in these specific areas (if it is considered acceptable to miss approximately 14 per cent of reports of randomized trials not indexed as randomized trials in MEDLINE for the period 1991 to date). Further research is required to assess the importance of the trials identified, and to assess how many trials, found either by handsearching or MEDLINE searching, were duplicated.

Further research is also required to determine the generalizability of these findings to other specialized and general health care journals published in the U.K. and to assess trends over shorter time periods (for example, annually). Research is also required to determine the comparative benefits of prospective handsearching and MEDLINE searching to identify reports of randomized trials in journals published outside the U.K. where language of publication may also be an important factor.
Finally, given the high number of reports of randomized trials currently published as abstracts or in supplements, and the importance of identifying these for systematic reviews, the U.S. National Library of Medicine should be encouraged to index all parts of a journal and not just articles, substantive editorials and letters.

ACKNOWLEDGEMENTS

We should like to acknowledge the work of the U.K. Cochrane Centre handsearchers in identifying reports of controlled trials. This work is part of a larger international project supported by the European Union BIOMED Programme (BMH4-CT98-3803). This study has been carried out as part of the programme of methodological research being undertaken at the U.K. Cochrane Centre, which is funded by the NHS R&D Programme.

DISCLAIMER

The views expressed in this paper represent those of the authors and are not necessarily the views or the official policy of the Cochrane Collaboration.

REFERENCES

13. McDonald S, Clarke M, Lefebvre C. Should we continue to handsearch MEDLINE-indexed journals prospectively?: some up-to-date empirical evidence. Poster presented at the 5th Annual Cochrane Colloquium, 8–12 October 1997; Amsterdam, The Netherlands.