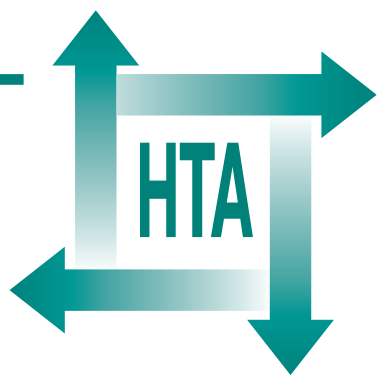
Health Technology Assessment 2003; Vol. 7: No. 34

Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system

P Royle
N Waugh

Health Technology Assessment
NHS R&D HTA Programme



How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system

P Royle*
N Waugh

Department of Public Health, University of Aberdeen, UK

* Corresponding author

Declared competing interests of authors: none

Published November 2003

This report should be referenced as follows:

Royle P, Waugh N. Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system. *Health Technol Assess* 2003;**7**(34).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE.

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

The research reported in this monograph was commissioned by the HTA Programme and funded as project number 02/30/01. Technology assessment reports are completed in a limited time to inform decisions in key areas by bringing together evidence on the use of the technology concerned.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

ISSN 1366-5278

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.
Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.

Abstract

Literature searching for clinical and cost-effectiveness studies used in health technology assessment reports carried out for the National Institute for Clinical Excellence appraisal system

P Royle* and N Waugh

Department of Public Health, University of Aberdeen, UK
* Corresponding author

Objective: To contribute to making searching for Technology Assessment Reports (TARs) more cost-effective by suggesting an optimum literature retrieval strategy.
Data sources: A sample of 20 recent TARs.
Review methods: All sources used to search for clinical and cost-effectiveness studies were recorded. In addition, all studies that were included in the clinical and cost-effectiveness sections of the TARs were identified, and their characteristics recorded, including author, journal, year, study design, study size and quality score. Each was also classified by publication type, and then checked to see whether it was indexed in the following databases: MEDLINE, EMBASE, and then either the Cochrane Controlled Trials Register (CCTR) for clinical effectiveness studies or the NHS Economic Evaluation Database (NHS EED) for the cost-effectiveness studies. Any study not found in at least one of these databases was checked to see whether it was indexed in the Science Citation Index (SCI) and BIOSIS, and the American Society of Clinical Oncology (ASCO) Online if a cancer review. Any studies still not found were checked to see whether they were in a number of additional databases.
Results: The median number of sources searched per TAR was 20, and the range was from 13 to 33 sources. Six sources (CCTR, DARE, EMBASE, MEDLINE, NHS EED and sponsor/industry submissions to National Institute for Clinical Excellence) were used in all reviews. After searching the MEDLINE, EMBASE and NHS EED databases, 87.3% of the clinical effectiveness studies and 94.8% of the cost-effectiveness studies were found, rising to 98.2% when SCI, BIOSIS and ASCO Online and 97.9% when SCI and ASCO Online, respectively, were added. The median number of sources searched for the 14 TARs that included an economic model was 9.0 per TAR. A sensitive search filter for identifying non-randomised controlled trials (RCT), constructed for MEDLINE and using the search terms from the bibliographic records in the included studies, retrieved only 85% of the known sample. Therefore, it is recommended that when searching for non-RCT studies a search is done for the intervention alone, and records are then scanned manually for those that look relevant.
Conclusions: Searching additional databases beyond the Cochrane Library (which includes CCTR, NHS EED and the HTA database), MEDLINE, EMBASE and SCI, plus BIOSIS limited to meeting abstracts only, was seldom found to be effective in retrieving additional studies for inclusion in the clinical and cost-effectiveness sections of TARs (apart from reviews of cancer therapies, where a search of the ASCO database is recommended). A more selective approach to database searching would suffice in most cases and would save resources, thereby making the TAR process more efficient. However, searching non-database sources (including submissions from manufacturers, recent meeting abstracts, contact with experts and checking reference lists) does appear to be a productive way of identifying further studies.

Contents

List of abbreviations

AMED	Allied and Complementary Medicine
ASCO	American Society of Clinical Oncology
BNF	British National Formulary
BNI	British Nursing Index
CABNAR	Commonwealth Agricultural Bureau Nutrition Abstracts and Reviews
CAM	complementary and alternative medicine
CCT	controlled clinical trial
CCTR	Cochrane Controlled Trials Register
CDSR	Cochrane Database of Systematic Reviews
CI	confidence interval
CRD	Centre for Reviews and Dissemination
DARE	Database of Abstracts of Reviews of Effectiveness
DEC	Development and Evaluation Committee
FDA	US Food and Drug Administration
HEED	Health Economics Evaluation Database
HMIC	Health Management Information Consortium
INAHTA	International Network of Agencies for Health Technology Assessment
IQR	interquartile range
ISTP	Index to Scientific and Technical Proceedings (now known as WoSP)
NDA	New Drug Application
NHS EED	NHS Economic Evaluation Database
NICE	National Institute for Clinical Excellence
NRR	National Research Register
QALY	quality-adjusted life-year
RCT	randomised controlled trial
SCI	Science Citation Index
SIGLE	System for Information on Grey Literature in Europe
SSCI	Social Sciences Citation Index
TAR	Technology Assessment Report
WoSP	Web of Science Proceedings

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or at the end of the table.

Executive summary

Background

In the UK, one part of the remit of the National Institute for Clinical Excellence (NICE) is to carry out a programme of technology appraisals. These are done to a fairly tight timetable in order not to delay the guidance on new technologies. Each appraisal is underpinned by a Technology Assessment Report (TAR) commissioned from a group of academic units.

As the TAR process is relatively new, and is still evolving, the methods used for its literature searching have been largely based on the well-established and documented methods used for Cochrane reviews. These involve comprehensive searching of a variety of sources to protect against bias, but can add substantially to the time and costs of carrying out a review.

However, resource constraints require that TARs are produced as efficiently as possible, and to a tight timetable, which means that not all of the Cochrane methods can be applied, or are appropriate. In addition, it is not known whether the marginal benefits of exhaustive searching justify the costs. The challenge for those undertaking TARs is to know how best to adapt and optimise, and extend when necessary, the Cochrane-based search strategies, so that searching can be done both rapidly and systematically.

Objective

To contribute to making searching for TARs more cost-effective by suggesting an optimum literature retrieval strategy, based on empirical data obtained from a sample of recent TARs, which balances comprehensiveness and efficiency.

Methods

A sample of 20 recent TARs was studied. All sources used to search for clinical and cost-effectiveness studies were recorded. In addition, all studies that were included in the clinical and cost-effectiveness sections of the TARs were identified, and their characteristics recorded, including author, journal, year, study design, study size and quality score. Each was also classified by publication type, and then checked to see whether it was indexed in the following databases: MEDLINE, EMBASE, and then either the Cochrane Controlled Trials Register (CCTR) for clinical effectiveness studies or the NHS Economic Evaluation Database (NHS EED) for the cost-effectiveness studies. Any study not found in at least one of these databases was checked to see whether it was indexed in the Science Citation Index (SCI) and BIOSIS, and the American Society of Clinical Oncology (ASCO) Online if a cancer review. Any studies still not found were investigated further to see whether they were in a number of additional databases.

Results

Sources searched

The median number of sources searched per TAR was 20, and the range was from 13 to 33 sources. Six sources (CCTR, DARE, EMBASE, MEDLINE, NHS EED and sponsor/industry submissions to NICE) were used in all reviews.

Clinical effectiveness studies

There were 424 studies in total. The publication types were: published 80%, meeting abstracts 11.3% and unpublished 8.7%. Eighty per cent of reviews included at least one abstract or unpublished study (60% included at least one abstract and 50% included at least one unpublished study). The median number of studies included per TAR was 19.5 (range 2–41). The median number of participants included per TAR was 2787 (range 69–97,570). Evidence from non-randomised controlled trial (RCT) studies was used in 45% of TARs. The proportion of studies classified either as published in full or as abstracts, and found indexed in the following databases, was: MEDLINE 82.7%, EMBASE 78.6% and CCTR 50.1%. The cumulative percentage of studies found after searching these three databases was 87.3%. Adding SCI, BIOSIS and ASCO Online increased this to 98.2%. Eighty-seven per cent of studies were indexed in both MEDLINE and EMBASE.

Cost-effectiveness studies

The 130 studies were classified as: published 73.1%, unpublished 23.8%, abstracts 1.5% and grey literature 1.5%. The median number of studies used was 4.0. The percentage of studies classified as either published in full or as abstracts, and found indexed in the following databases, was: MEDLINE 86.6%, EMBASE 86.6% and NHS EED 40.2%. The cumulative percentage of these studies found indexed after searching the three databases was 94.8%. Adding SCI and ASCO Online increased this to 97.9%.

Studies used in the economic modelling

The 121 articles were classified as: published 50.4%, abstracts 5.0%, reference sources 17.4%, unpublished 17.4% and grey literature 9.8%. The median number of studies used for the 14 TARs that included an economic model was 9.0 per TAR.

Search terms for identifying non-RCTs

A sensitive search filter, constructed for MEDLINE and using the search terms from the bibliographic records in the included studies, retrieved only 85% of the known sample. Therefore, it is recommended that when searching for non-RCT studies a search is done for the intervention alone, and records are then scanned manually for those that look relevant.

Conclusions

Searching additional databases beyond the Cochrane Library (which includes CCTR, NHS EED and the HTA database), MEDLINE, EMBASE and SCI, plus BIOSIS limited to meeting abstracts only, is seldom effective in retrieving additional studies for inclusion in the clinical and cost-effectiveness sections of TARs (apart from reviews of cancer therapies, where a search of the ASCO database is recommended). A more selective approach to database searching would suffice in most cases and would save resources, thereby making the TAR process more efficient. However, searching non-database sources (including submissions from manufacturers, recent meeting abstracts, contact with experts and checking reference lists) does appear to be a productive way of identifying further studies.

Chapter 1
Introduction

Health technology assessment aims to answer a series of questions:

- Does it work?
- At what cost?
- Is it worth it?

These questions may contain subsidiary questions. For example, the 'does it work' question may have subsidiary questions about for whom, at what dose, and so on.

The International Network of Agencies for Health Technology Assessment (INAHTA) (see http://www.inahta.org) defines healthcare technology as:

> "prevention and rehabilitation, vaccines, pharmaceuticals, devices, medical and surgical procedures, and the systems within which health is protected and maintained."

INAHTA then defines health technology assessment as:

> "A multidisciplinary field of policy analysis which studies the medical, social, ethical, and economic implications of development, diffusion, and use of health technology."

In the UK, one part of the remit of the National Institute for Clinical Excellence (NICE) is to carry out a programme of technology appraisals. These are done to a fairly tight timetable in order not to delay the guidance on new technologies. Each appraisal is underpinned by a Technology Assessment Report (TAR) commissioned from a group of academic units.

TARs have a number of components, the two main ones being a systematic review of the evidence on clinical effectiveness of the technology, and an economic evaluation of relative cost-effectiveness. Other sections may include an introduction to the disease (frequency, natural history, effects on those who suffer from it) and its current alternative treatments, including how successful they are.

The purpose of a review of the literature is to bring together and summarise the evidence. The need for reviews has become greater over time owing to the increasing number of scientific publications. For example, the MEDLINE database currently indexes about 5000 titles and contains over 11.6 million records, and its rate of growth is steadily increasing; in the past 10 years (1993–2002) approximately 4.4 million records were added to the database, whereas in the previous decade the comparative number was 3.5 million, and in the decade before that it was 2.6 million.

Traditional narrative reviews have been recognised as being unsystematic and prone to bias.[1] They are usually written by experts in the field, and use informal and subjective methods to collect and interpret information. The main advantage of narrative reviews is that they require less time and fewer resources to prepare than systematic reviews. However, they are not based on a comprehensive search for studies, and there is no detailed description of the quality of the included studies. They are therefore susceptible to bias (e.g. in favour of a technology). The need for properly systematic reviews is therefore recognised.[1]

The evolution of systematic reviews

A systematic review has been defined as a review of a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies.[2] Systematic reviews should also include an assessment of the quality of the included studies.

There are two major advantages of systematic reviews (or meta-analyses). First, by combining data they improve the ability to study the consistency of results (i.e. they give increased power). Second, similar effects across a wide variety of settings and designs provide evidence of robustness and transferability of results to other settings. If the studies are inconsistent between settings then the sources of variation can be examined.[3]

The main force behind the development of systematic review methodologies has been the Cochrane Collaboration,[4] which started in 1992 with the opening of the UK Cochrane Centre and has developed rapidly into an international collaboration. It has established various methodological groups. The history of the Cochrane Collaboration has been recorded in various places[5] and need not be repeated in detail here. One result of the methodological work has been the rising quality of Cochrane reviews, which have earned a reputation for being of higher quality and hence more reliable than other reviews.[6]

The evolution of health technology assessment in the UK

NICE was set up as a Special Health Authority for England and Wales in 1999.[7] It is part of the NHS, and its role is to provide patients, health professionals and the public with authoritative, robust and reliable guidance on current 'best practice'. The guidance covers both individual health technologies (including medicines, medical devices, diagnostic techniques and procedures) and the clinical management of specific conditions. The remit of NICE is to examine the effectiveness, cost-effectiveness and broader impact of technologies.

Since the establishment of NICE, we have seen a sequence of increasing sophistication of health technology assessment 'rapid reviews' in the UK. These began with those produced by the Development and Evaluation Committees (DECs)[8] (which had themselves been evolving), followed by the initial reviews for NICE, and then by a process of incremental developments to the current TARs for NICE. The term 'rapid reviews' is used to distinguish the TARs from the 'exhaustive reviews' which had been commissioned by the Centre for Reviews and Dissemination (CRD) or by the Health Technology Assessment (HTA) programme. TARs are produced to a tight timescale and with a fixed budget, whereas the more traditional systematic health technology assessment reviews have more resources and two to three times the timescale.

One of the developing features in technology assessments, particularly those done for NICE, is the increase in economic content. This is probably largely because technologies referred to NICE usually have some evidence of clinical effectiveness (typically the trial data required for drug licensing) and hence decisions on guidance often depend more on cost-effectiveness. TARs therefore require a wider range of evidence than reviews focusing purely on effectiveness, such as Cochrane reviews.

TARs are different from Cochrane reviews in a number of ways. Some of these are:

- **Scope and nature of the questions asked:** Cochrane reviews usually ask a much more focused question than TARs, which tend to ask broader questions, and can have a number of associated questions or comparisons. In addition, Cochrane reviews usually aim to determine simply whether one therapy is better than another, whereas TARs will ask how much better the therapy is, and at what cost. This is partly because TARs need to be set in a policy context of opportunity cost. Cost-effectiveness is usually expressed via the common currency of cost per quality-adjusted life-year (QALY), which enables the value of very different types of healthcare interventions to be compared. For example, QALYs allow comparison of the relative benefits of reductions in mortality (life-years gained) and improvements in quality of life (but without survival gains).
- **Study designs used:** Cochrane reviews are mainly (although not exclusively) based on randomised controlled trials (RCTs), whereas TARs are more likely to include other study designs. This is partly because TARs address a wider range of questions, but also because they cover topics for which there are no RCTs.
- **Inclusion of economic evaluations and costs:** Cochrane reviews do not usually assess the cost-effectiveness of interventions, whereas TARs do, since the assessment teams are asked to produce a cost per QALY bottom line whenever possible.
- **Time-frames:** TARs are produced to a tight and rigid deadline, whereas Cochrane reviews tend to have a more flexible timetable.
- **Authors:** Cochrane reviews tend to be done by volunteers with a professional subject interest, whereas TARs are commissioned by the Department of Health, which allocates them to the various TAR teams. The reviewers on the TAR teams are experts in review methodology, and independent of vested interests, but do not usually have detailed knowledge of the disease topic under review. Different subject specialists and experts are recruited as advisors for the TARs as needed, on a report-by-report basis.
- **Industry input:** TARs have access to invited submissions from relevant industry groups and manufacturers. As there are often large financial

interests at stake in the outcome of the NICE guidance, such groups will often expend considerable resources in preparing their submissions. The version of the TAR made available to NICE may include commercially confidential information supplied by industry, such as unpublished data and economic models. The confidential information is removed from the version of the TAR that is made available in the public domain. By contrast, Cochrane reviews are less likely to have unpublished studies, although some reviewers do approach known researchers or industry for details of published studies.

- **Impact on policy:** TARs directly inform health policy, as is it now mandatory for strategic health authorities and primary care trusts in England to follow NICE guidance. Although Cochrane reviews can provide the raw material for TARS, they do not usually directly impact on policy in the same way. As the NICE process and products receive great attention from industry, professional and consumer groups, this exposure puts additional pressure on the authors of TARs to produce reports that can withstand criticism.

Hence, it can be seen that various factors have led to the TARs moving from being rapid reviews to being rapidly produced systematic reviews and economic evaluations. They have to be based on a comprehensive review of the literature, in the sense of identifying all of the important studies. For assessing effectiveness, these are usually RCTs, but can include other study designs if there is insufficient evidence from RCTs. TARs will be criticised if an RCT, or other relevant evidence, is missed.

Rationale

The challenge for those doing literature searches for TARs is to decide on the best use of limited resources in the face of two conflicting pressures. On the one hand, time and cost constraints can preclude exhaustive searching, such as that recommended for Cochrane reviews. On the other hand, one needs to protect against bias and random error by doing comprehensive searches, to ensure that as much as possible of the relevant evidence is identified. At the same time, one needs to be mindful of the diminishing returns, in both quality and quantity of evidence, obtained from extended searching beyond a particular point.

The overall aim of this study is to contribute to making searching for TARs more cost-effective by suggesting an optimum literature retrieval strategy, based on empirical data obtained from a sample of recent TARs, which balances comprehensiveness and efficiency.

Objectives

The main objectives are:

- To survey the frequency with which the different sources have been searched in recent TARs for NICE.
- To measure the proportion of trials (both individually and cumulatively) cited in the clinical and cost-effectiveness sections and indexed in the major databases, including the Cochrane Library, MEDLINE, EMBASE, Science Citation Index (SCI) and BIOSIS (it is assumed that most TARs will rely mainly on RCTs for the clinical effectiveness data).
- To determine which other sources (i.e. those found outside the major databases) indexed studies that were included in TARs, and to consider the characteristics of those studies indexed outside the major databases.
- To analyse the terms used to describe the study design within the full bibliographical records of the studies that are found. This should provide information on the proportion of records that describe the study design in the title, abstract or indexing fields, and hence on which terms to use in search filters for study designs.

Chapter 2

Literature review

The limitations of systematic reviews, and the importance of literature searching and trial quality, have been comprehensively reviewed by Egger and colleagues,[9,10] and therefore will not be repeated here. The following literature review will focus on articles that have been published since the aforementioned reviews and are relevant to this study.

A search was done of the databases MEDLINE, EMBASE, the Cochrane Methodology Register, and the Cochrane Database of Methodology Reviews (CDSR), and limited to the publication dates 2000–2003 and to English language studies.

Background

The distribution and accessibility of trials in the medical literature span a wide spectrum. At one end there are trials that are either inaccessible (e.g. confidential industry data) or difficult to access (e.g. those in grey literature and journals not indexed in MEDLINE); at the other end there are those that are readily accessible (e.g. articles in high-impact English language journals indexed in MEDLINE). Consequently, it is usually easy to identify and access some relevant trials, but progressively harder to find more.

To ensure that no evidence is missed, one can carry out searches of many sources, including a wide variety of databases, include non-database sources (e.g. handsearching journals and conference abstracts), and contact manufacturers and experts in the area. However, at the same time, resource constraints also require that the TARs are produced as efficiently as possible. As the TAR process is relatively new and is still evolving, the methods used for its literature searching have been largely based on the well-established and documented methods used for Cochrane reviews. The Cochrane Handbook[2] recommends comprehensive searching of a variety of sources, and using a systematic approach to select studies for inclusion in the review to protect against bias. However, given the time and resource constraints that apply to TARs, the Cochrane methods cannot all be applied. For example, there is rarely time or resources to include papers in languages other than English, or to contact authors for unpublished material.

Publication bias

One of the aims of systematic reviews is to avoid bias of different kinds, one of the major ones being publication bias.[11] This is a bias in the published literature, where the publication of research depends on the nature and direction of the study results.[2] Studies in which an intervention is not found to be effective are sometimes not published. This bias is a major threat to the validity of systematic reviews, as those that fail to include unpublished studies may overestimate the true effect of an intervention. Therefore, to minimise publication bias, exhaustive searches are thought to be necessary when doing systematic reviews in order to maximise the retrieval of relevant studies.

Various additional types of bias can be introduced into reviews. These have been comprehensively reviewed by Egger and colleagues.[12] Some recent studies on biases of particular relevance to this study are outlined below.

Location bias

There is evidence that papers that contain novel scientific ideas and unexpected observations, especially those that are not consistent with the prevailing paradigm in the field, can have difficulty being published.[13,14] They may be less well accepted by journals and peer reviewers, and therefore appear in less prestigious, low-circulation journals, which are not indexed by major databases. Conversely, it is sometimes the case that new and exciting results, especially from trials that show large treatment effects, are more likely to be published in high-impact journals, as they are deemed to be more newsworthy and hence attract publicity and citations to the journal. Trial results considered to be less 'interesting' may be published in lower impact or local journals that may not be widely disseminated, and hence less likely to be indexed in the major databases. Location bias affects the probability of articles

being identified when searching for studies for inclusion in systematic reviews.

A study was done to investigate systematically location bias in controlled clinical trials (CCTs) in the field of complementary and alternative medicine (CAM).[15] Trials were categorised by whether they appeared in CAM journals or mainstream medical journals, and by their direction of outcome, methodological quality and sample size. A predominance of positive trials was seen in non-impact factor CAM and mainstream medical journals. In high-impact mainstream medical journals there were equal numbers of positive and negative trials. Quality scores were significantly lower for positive than for negative trials in non-impact factor CAM journals. There were no significant differences between quality scores of positive and negative trials published in mainstream medical journals, except for high impact factor journals, in which positive trials had significantly lower scores than negative trials. It was concluded that the location of trials, in terms of journal type and impact factor, should be taken into account when the literature on complementary therapies is being examined.

Two of the objectives of a recent report by Egger and colleagues[9] were: (1) to examine the characteristics of clinical trials that are difficult to locate, and also trials that are of lower quality; and (2) to compare within meta-analyses the treatment effects reported in trials that are difficult to locate with trials that are more accessible, and of trials of lower quality with those of higher quality. They defined 'difficult to locate' trials as unpublished trials, trials published in languages other than English and trials published in journals not indexed in the MEDLINE database.

They included meta-analyses from four sources: high-impact medical journals, the CDSR, the Database of Abstracts of Reviews of Effectiveness (DARE) and HTA reports. Only meta-analyses based on comprehensive literature searches, and which combined the binary outcomes of at least five CCTs, were selected.

Within each meta-analysis, pooled effect estimates were calculated separately for the trials that were difficult to locate and the remaining trials, applying the same statistical model used by the original authors. For each meta-analysis, a ratio of the pooled estimates was derived; the percentage change in the pooled effect estimate that occurred when trials that are difficult to locate were excluded was calculated.

They found that the importance of trials that are difficult to locate appeared to vary across medical specialities, with a large proportion of such trials coming from complementary medicine. Trials that were difficult to locate also tended to be smaller and of lower methodological quality than trials that were easily accessible and published in English. Poor-quality trials showed more beneficial effects than good-quality trials.

Their findings led them to conclude that rather than preventing bias through extensive literature searches, bias could be introduced by including trials of low methodological quality. They believe that in situations where resources are limited, thorough quality assessments should take precedence over extensive literature searches and translations of articles.

The view that the more obscure journals tend to include the poorer quality articles is supported by the work of Lee and colleagues.[16] They found that high citation rates, impact factors and circulation rates, and low manuscript acceptance rates and indexing on the Brandon/Hill Library List (a core list of journals recommended for medical libraries)[17] appear to be predictive of higher methodological quality scores for clinical research articles.

Impact of grey literature on systematic reviews

The term 'grey literature' is used here to include literature that has not been formally published in peer-reviewed journals. It includes conference proceedings, meeting abstracts, research reports, book chapters, unpublished data, dissertations, policy documents and personal communications.

A recent systematic review looked at research studies that have investigated the impact of grey literature in meta-analyses of randomised trials of healthcare interventions.[18] Eight studies, covering a variety of different areas of healthcare, met the inclusion criteria. Four studies contained multiple meta-analyses, which compared the inclusion and exclusion of grey literature on the pooled effect estimate of the meta-analyses. All four found that published trials showed an overall greater treatment effect than grey trials, but this difference was statistically significant in only two of the four studies.

The other four studies contained single meta-analyses, where the difference between the

treatment effect of grey and published trials was explored in a sensitivity analysis. None of these studies found a statistically significant difference, although three found that published trials showed an overall greater treatment effect than grey trials, whereas one study found that published trials showed no effect of treatment and that grey trials showed a negative treatment effect. The two studies that assessed the methodological quality of the included trials found that the published trials were of higher quality than the grey trials. For all eight included studies the most common type of grey literature was abstracts (49%), followed by unpublished data (33%).

The conclusions were that published trials are generally larger and may show an overall greater treatment effect than grey literature trials, and therefore people conducting systematic reviews need to ensure that they search for trials in the grey, as well as published, literature in order to minimise bias in their review. It was noted that this may have particular implications in meta-analyses containing only a few trials, where the impact of excluding grey trials has the greatest potential to introduce bias.

A review of the rate at which results in abstracts are subsequently published in full, and the time interval between presentation at a meeting and full publication, found that only 45% of all studies first presented as abstracts were published in full within 2 years following presentation at meetings or publication as a summary report.[19] In addition, full publication of studies initially appearing as abstracts appeared to be biased, in that significant results were published more frequently than non-significant results. This suggests that systematic reviews that exclude searches for meeting abstracts are in danger of missing a significant portion of the relevant data and of overestimating treatment effects.

Impact of unpublished data reported in US Food and Drug Administration (FDA) reviews on systematic reviews

The impact of unpublished data reported in FDA reviews of New Drug Applications (NDAs) has been investigated in a systematic review.[20] The review found no meaningful difference between the methodological quality of published trials (most of which were manufacturer sponsored) and the manufacturer-sponsored unpublished trials included in the FDA reviews. The authors suggest that the inclusion of FDA data in systematic reviews and meta-analyses should be considered in situations when: (1) there is a paucity of published data, and (2) there is an a priori reason to suspect that FDA data may be systematically different to published data.

However, it is not universally agreed that unpublished data should be included in systematic reviews.[21] Some do not favour their inclusion on the basis that they have not been peer reviewed, and are therefore less reliable than published data. However, a recent systematic review concluded that at present there is little empirical evidence to support the use of editorial peer review as a mechanism to ensure quality of biomedical research, despite its widespread use and costs.[22]

The effectiveness of extended search strategies

Savoie and colleagues[23] undertook a prospective analysis of extended search methods (specialised databases or trial registries, reference lists, handsearching, personal communication and the Internet) for identifying RCTs for two systematic reviews, one on acupuncture in the treatment of addiction and the other on lipid lowering. They searched four major databases and conducted additional database searches to improve comprehensiveness.

It was found that the extended searching identified 94 additional RCTs for the systematic reviews; as a proportion of all RCTs included, this was 42.9% (9) for the acupuncture project and 23.5% (85) for the lipid-lowering project. This extra retrieval for the latter project was largely attributable to the search of the Cochrane Library. However, a post hoc analysis 1 year later showed that 75 of the 94 additional RCTs were now indexed in the major databases. It was thought that this was because items identified through conference abstracts or personal communication had subsequently been published.

However, the authors did not assess the quality of the extra RCTs retrieved or their impact on the results of systematic reviews, and acknowledged that this information is needed before one can definitively determine the value of extended search methods.

A study by Sampson and colleagues aimed to identify sources of reported RCTs in the field of paediatric CAM.[24] Reports of 908 RCTs were identified by searching MEDLINE and 12 additional bibliographic databases, and by reviewing the reference lists of previously identified paediatric CAM systematic reviews. It was found that a search of MEDLINE alone could potentially identify 97.7% of these reports, and a search combining MEDLINE, EMBASE and CAB Health could achieve 99.4% retrieval.

Avenell and colleagues evaluated a search plan for identifying RCTs for a Cochrane review on nutritional supplementation after hip fracture.[25] They identified 15 RCTs; 11 by database searching, two unpublished trials via experts in the field and one conference abstract from handsearching. A subsidiary analysis examined the degree to which the available evidence would be deficient if only one of the six electronic databases had been used. Of the 1054 participants recruited into the 15 trials, only 58% would have been included if EMBASE only were used, 56% if MEDLINE or HEALTHSTAR only were used, 48% if Commonwealth Agricultural Bureau Nutrition Abstracts and Reviews (CABNAR) only were used, 42% if BIOSIS only were used, and 0% if CINAHL only were used.

Thus, they recommended that reviews in nutrition include searches of several electronic databases, including the Cochrane Controlled Trials Register (CCTR), in conjunction with handsearching of journals (particularly recent issues and those that contain conference proceedings), consultation with experts and searching the reference lists of trial reports.

A study was done to compare the performance of MEDLINE and EMBASE for the identification of CCTs evaluating the management of selected musculoskeletal diseases.[26] Selected journals were also handsearched to identify CCTs not retrieved by either database. Of 243 papers about CCTs, two-thirds were retrieved by both databases and one-third by only one database. An additional 16 CCTs not retrieved by either database were identified through handsearching. No significant differences were observed in the mean quality scores and sample size of the CCTs missed by MEDLINE compared with those missed by EMBASE. The reasons why the various CCTs were not retrieved were investigated and it was found that for MEDLINE, in general, it was because the specific journal was not indexed or the subject heading terms assigned were inadequate. For EMBASE, references were most frequently missed owing to the assignment of indexing terms.

It was therefore concluded that the use of MEDLINE alone to identify CCTs is inadequate, and the use of two or more databases and handsearching of selected journals are needed to perform a comprehensive search.

An analysis was performed of the sources searched in Cochrane reviews, which were new to the 2001, Issue 1, of the Cochrane Library, and included only RCTs.[27] The proportion of trials that were indexed in the major databases was determined, and their quality compared with the trials found from other sources. Extended database searching beyond the major databases (Cochrane, MEDLINE and EMBASE) retrieved only a very small percentage of extra trials, and these were generally of poorer quality than those trials that were easily found. Contacting authors and manufacturers to find unpublished trials appeared to be a more effective method of obtaining the additional better quality trials.

A study on the efficiency of searching the grey literature in the field of palliative care failed to confirm the usefulness of grey literature searching.[28] A systematic review into palliative care team effectiveness was undertaken, and addressed the question of whether grey literature searching was worth the time and money spent. The authors conducted main database searches and then augmented them with a grey literature search. They estimated that they spent approximately 300 hours on contacting the experts in the area, and 2–3 hours on searching the System for Information on Grey Literature in Europe (SIGLE) database. However, this comprehensive search was unsuccessful in obtaining unpublished studies as it did not add to the recall of the overall search strategy.

Therefore, they concluded that, for now, grey literature searching represents an unjustifiable use of financially constrained resources when conducting a systematic review in the field of palliative care. They speculated on reasons for the failure to obtain any studies from their comprehensive grey literature search. These included the fact that evaluative research in palliative care is an emerging field, and it is therefore likely that indexing in databases is of variable quality. It is equally possible that in this rapidly developing field, researchers are protective of their data and are only likely to disseminate them through peer-reviewed channels.

Conclusion

There appears to be some empirical evidence of the marginal benefits of exhaustive searching in terms of the quantity of studies identified, as it would seem that in many cases extended searching does identify extra studies. However, it is difficult to generalise between topics, as the exact degree of the effectiveness of extended searching appears to be quite heavily subject dependent. It is also likely that it in some disciplines that are not yet well developed, the effectiveness of exhaustive searching could change over time as the field matures.

However, there is still very little evidence as to whether extended searching of additional databases and grey literature is cost-effective. To determine this, one would need to determine: (1) the total costs involved in identifying and obtaining the difficult to find studies from the additional searching; (2) their relative quality compared with the more easily found studies; and (3) how much difference the inclusion of the extra studies found makes to the final results of a systematic review.

Chapter 3

Methods

Selection of the sample of TARs to be used in this study

Initially, the titles of all the TARs published in the past 2 years (as of December 2002) were downloaded from the NICE web site. The reports were classified into six different categories by both authors:

1. Drug
2. Therapeutic device
3. Surgical procedure
4. Setting
5. Therapeutic procedure (non-surgical)
6. Diagnostic.

It was anticipated that there would be a preponderance of drug reports, so to obtain a sample representative of the range of TARs, 20 titles were chosen so as to include some that were assessments of non-pharmacological interventions, in case the optimum search strategy differed. Each TAR title was assigned a unique R (Review) number.

Sources used to search for clinical and cost-effectiveness studies

All sources used to identify clinical effectiveness studies in 20 selected TARs were scrutinised and the data listed in a spreadsheet. Also noted were: the date on which the searching was done, whether any additional searches were done, any restrictions on the searching (i.e. date, language, format and study design), whether search filters were used, and the year and issue number of the version of the Cochrane Library searched. If a report stated that the Cochrane Library had been searched (without specifying which sections searched) then it was assumed that all sections of the Library were searched. [The Cochrane Library comprises several databases, including the CDSR, CCTR, DARE, the Health Technology Assessment Database and the NHS Economic Evaluation Database (NHS, EED).]

Recording details of clinical effectiveness studies

Those studies that were included in the clinical effectiveness section (i.e. all studies that were data extracted) were identified from the TAR and their details listed in a spreadsheet. Each study was given a unique number. The details recorded for each study (where the information was available) were:

- first author
- year of publication
- journal title
- study design: this was exactly as described in the TARs. No attempt was made to standardise the descriptions. The full papers were only examined for those articles described as RCTs but not indexed in CCTR
- publication type: each study was classified into one of three possible categories: published, abstract or unpublished. These categories were decided on the basis of a pilot study. 'Published' referred to anything that had been published in full in a peer-reviewed journal or series, 'abstract' referred to meeting abstracts and 'unpublished' was anything else that did not fall into the other two categories
- sample size or number of participants: the information was sought from the data extraction sheet in the TAR. If the data were not available from the TAR, then the information was obtained from published study where possible
- quality score of study: where available, the data were taken from the TAR, unless otherwise stated
- journal impact factor: taken from the ISI *Journal Citation Reports* 2001.

Each study was then checked to see whether it was indexed in all of the following databases:

- The Cochrane Controlled Trials Register (CCTR)* on the Cochrane Library. The version of the Cochrane Library used for the searching for the TAR was determined from the methods section of the TAR, and this version was searched. The current version of the Cochrane

Library (2002, Issue 4) was also searched to see whether the study was present.
- MEDLINE.
- EMBASE.

*The CCTR has recently been renamed the Cochrane Central Register of Controlled Trials (CENTRAL). However, it will be referred to in this report as CCTR, as this was its name in the versions searched in this study.

If any study was classified as either published or abstract, and was not found in any of the databases above, it was checked to see whether it was indexed in SCI or BIOSIS. Any studies still not found in any databases were checked to see whether they were in any of the following databases (the specific databases chosen depended on the subject area and type of publication of each article): American Society of Clinical Oncology (ASCO) Abstracts database, PsycINFO, Web of Science Proceedings (WoSP), National Research Register (NRR), Zetoc, Health Management Information Consortium (HMIC), Conference Papers Index, Dissertation Abstracts Online, Index to Theses, CINAHL, British Nursing Index (BNI) or a cited reference search using SCI.

The full bibliographic records of the studies found indexed in any database were downloaded into Reference Manager version 9.5. If the study was found in MEDLINE, then the MEDLINE record was downloaded in preference to that from any of the other databases. If the study was not in MEDLINE, then the CCTR or EMBASE records were used if available. If a study was not found indexed anywhere, then the details available from the bibliography of the TAR were entered manually into Reference Manager.

Recording details of cost-effectiveness studies

This was done in two parts to reflect the usual content of the economics section, which usually consists of a review of published economic studies and the construction of an economic model by the TAR team. The modelling process requires cost and other data.

Studies used in the economic evaluation

Those studies that were included in economic evaluation sections were identified from the TAR and their details listed in a spreadsheet. The details recorded were:

- first author
- publication year
- journal title
- publication type: each study was classified into one of four possible categories: published, abstract, grey literature or unpublished. These categories were decided on the basis of a pilot study.

Each study was then checked to see whether it was indexed in:

- NHS EED: the version of the Cochrane Library used for the searching for the TAR was determined from the methods section of the TAR, and this version was searched. The current version of the Cochrane Library (2002, Issue 4) was also searched to see whether the study was present.
- MEDLINE.
- EMBASE.

If any study was not found in these databases additional searches of SCI, BIOSIS, ASCO Online or a cited reference search using SCI were carried out until the study was found.

Recording the terms used to describe the study as being an economic evaluation

Any potential search terms that could be used to identify the study as being an economic evaluation were checked for in the bibliographic record in the fields: Title, Abstract, Subject Heading field. Any terms thought relevant according to the judgement of the first author, and the field in which they occurred, were recorded in a spreadsheet.

Sources used for the economic modelling

Those studies that were cited in the section on economic modelling were identified from the TAR and their details listed in a spreadsheet. The details recorded were:

- first author
- publication year
- name of source
- publication type: each study was classified into one of five possible categories: published, abstract, grey literature, reference source or unpublished. These categories were decided on the basis of a pilot study.

Determining the 'age' of the included studies

This was defined as the difference in the number

of years between the date on which the search was done and the year of publication of the article. Any study that was published in the same year as the search was done was recorded as 0 years.

Additional data noted for each report

Also noted for each TAR were the following: the methods used to quality assess the studies; whether separate searches were done for clinical and cost-effectiveness studies or whether they were combined, and whether any additional searches were done (apart from background, clinical and cost-effectiveness).

Analyses

The data recorded in the Excel spreadsheet were then exported into the MS Access database to run queries. Statistical analyses were done using the program SPSS for Windows version 10.

Chapter 4
Results

Selection of the study sample

The titles of the 27 TARs published in the 2 years between December 2000 and 2002 were downloaded from the NICE website. Each title was classified into one of the six categories (see Chapter 3) by both authors. Twenty-one (78%) were classified as 'drug'. There were two discrepancies in classification that were resolved by discussion.

Twenty TARs were chosen for analysis. All six of the non-drug reports were selected. The remaining 14 TARs, classified as 'drug', were chosen to give a representative sample of different interventions and diseases; for example, if one had already been chosen on an arthritis drug or breast cancer, a report on a different type of drug or disease was selected in preference.

Table 1 shows the full titles of the 20 TARs finally selected, their NICE publication number and the review (R) numbers and classifications assigned to them for this study.

A summary of the classifications of the 20 TARs selected for this study were: drug = 14, devices = 2, surgery/device = 1, surgery = 1, therapeutic procedure = 1 and setting = 1. It can thus be seen that drug interventions predominated; the most common were on drugs for cancer, of which there were seven.

A survey of current practice in searching for TARs

In total, 67 different sources were used over the 20 TARs. The complete list of sources searched is shown in the table in the Appendix. The numbers of sources used to search for both clinical and cost-effectiveness studies were combined. Twelve reviews listed their searches for clinical and cost effectiveness separately, and the remaining eight appeared to do both searches together.

Forty-eight of the 67 sources used were electronic databases, with the remainder being non-database sources, including sponsor or industry submissions, consulting reference lists, web searching, and contact with experts and manufacturers.

There appears to be a skewed distribution of sources searched: the first 15 (22.4%) of the sources were used in more than half of the reviews, but there is a long tail of sources only searched in a few reports, with nine (13%) sources being searched in only two reviews, and 17 (25%) searched in only one review.

Table 2 shows the ten most commonly used sources. It can been seen that all of the first six sources were searched in all 20 TARs, that CDSR and NRR were each searched in 18 TARs, and the HTA database and SCI in 17 and 16 TARs, respectively. A further analysis of the data revealed that 15 of the TARs searched at least all of these ten sources.

The combined numbers for clinical and cost-effectiveness studies searched per review are shown in *Table 3*. The data show that there is quite a large variation between reviews in the number of sources searched, from a minimum of 13 for metal-on-metal hip resurfacing arthroplasty for the treatment of hip disease, to a maximum of 33 for the review of computerised cognitive behaviour therapy for depression and anxiety.

The median number of sources searched per review was 20, the mean 21 and the interquartile range (IQR) from 16 to 27.

Restrictions on searching by language

- Eight reviews stated that they did not include non-English studies (four of these eight reviews stated that they searched for the non-English papers, but did not include them in the review).
- Seven reviews did not mention whether there were any restrictions on language when searching.
- Four reviews specifically mentioned that language restrictions were not used.
- One review stated that only studies in English, French, Dutch or German were considered for inclusion in the review.

Therefore, 40% of reviews explicitly stated that they included English language studies only, 35% of reviews did not mention whether there were any restrictions on the language when searching, 20% specified that there were no restrictions and 5% mentioned partial restrictions.

TABLE 1 *Titles of TARs selected for analysis*

Review no.	NICE ref. no.	Full title	Classifications
R1	26	A rapid and systematic review of the clinical effectiveness and cost-effectiveness of paclitaxel, docetaxel, gemcitabine and vinorelbine in non-small-cell lung cancer	Drug
R2	38	The clinical effectiveness and cost-effectiveness of inhaler devices used in the routine management of chronic asthma in older children: a systematic review and economic evaluation	Device
R3	44	A systematic review of the effectiveness and cost-effectiveness of metal-on-metal hip resurfacing arthroplasty for treatment of hip disease	Surgical procedure/ device
R4	39	The clinical effectiveness and cost-effectiveness of bupropion and nicotine replacement therapy for smoking cessation: a systematic review and economic evaluation	Drug
R5	33	A rapid and systematic review of the evidence for the clinical effectiveness and cost-effectiveness of irinotecan, oxaliplatin and raltitrexed for the treatment of advanced colorectal cancer	Drug
R6	28	A rapid and systematic review of the clinical effectiveness and cost-effectiveness of topotecan for ovarian cancer	Drug
R7	29	Fludarabine as second-line therapy for B-cell lymphocytic leukaemia: a technology assessment.	Drug
R8	37	Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma: a systematic review and economic evaluation	Drug
R9	31	The clinical effectiveness and cost-effectiveness of sibutramine in the management of obesity: a technology assessment	Drug
R10	46	The clinical effectiveness and cost-effectiveness of surgery for people with morbid obesity: a systematic review and economic evaluation	Surgical procedure
R11	34	The clinical effectiveness of trastuzumab for breast cancer: a systematic review	Drug
R12	35	A systematic review of effectiveness and economic evaluation of new drug treatments for juvenile idiopathic arthritis: etanercept	Drug
R13	42	Clinical effectiveness and cost-effectiveness of growth hormone in children: a systematic review and economic evaluation	Drug
R14	36	The effectiveness of infliximab and etanercept for the treatment of rheumatoid arthritis: a systematic review and economic evaluation	Drug
R15	45	A systematic review and economic evaluation of pegylated liposomal doxorubicin hydrochloride for ovarian cancer	Drug
R16	51	A systematic review and economic evaluation of computerised cognitive behaviour therapy for depression and anxiety	Therapeutic procedure
R17	48	Systematic review of the effectiveness and cost-effectiveness of home versus hospital or satellite unit haemodialysis for people with end stage renal failure	Setting
R18	49	The effectiveness and cost-effectiveness of ultrasound locating devices for central venous access	Device
R19	41	A review of the clinical effectiveness and cost-effectiveness of routine anti-D prophylaxis for pregnant women who are rhesus-negative	Drug
R20	47	A systematic review update of the clinical effectiveness and cost-effectiveness of glycoprotein IIb/IIIa antagonists	Drug

TABLE 2 *The ten most commonly used sources for searching in TARs*

	Name of source	No. of reviews in which source was searched
1	CCTR	20
2	DARE	20
3	EMBASE	20
4	MEDLINE	20
5	NHS EED	20
6	Sponsor/industry submissions to NICE	20
7	CDSR	18
8	NRR	18
9	HTA database	17
10	SCI	16

Restrictions by publication type

- Four reviews stated that they excluded items published as abstracts only.
- One review excluded abstracts for the economic evaluation searching, but not the clinical effectiveness searches.

Restrictions by date

- Three reviews explicitly mentioned that there were no restrictions on the date of searching.
- One review mentioned some restrictions on searching; this was an update of an earlier review. For the other reviews it was assumed, unless otherwise stated, that databases were searched from their inception.

Use of search filters

- Fourteen reviews mentioned the use of a methodological search filter when searching databases for clinical effectiveness studies. The most commonly used was a sensitive search filter for RCTs.

TABLE 3 *Number of sources searched per review for clinical and cost-effectiveness studies*

Review no.	Short title	No. of sources searched
R16	Computerised cognitive behaviour therapy for depression and anxiety	33
R2	Inhaler devices used in the routine management of chronic asthma in older children	29
R18	Ultrasound locating devices for central venous access	28
R4	Bupropion and nicotine replacement therapy for smoking cessation	28
R15	Pegylated liposomal doxorubicin hydrochloride for ovarian cancer	27
R9	Sibutramine in the management of obesity	25
R17	Home versus hospital or satellite unit haemodialysis for people with end stage renal failure	24
R13	Growth hormone in children	24
R10	Surgery for people with morbid obesity	23
R1	Palcitaxel, docetaxel, gemcitabine and vinorelbine in lung cancer	20
R19	Routine anti-D prophylaxis for pregnant women who are rhesus-negative	20
R5	Irinotecan, oxaliplatin and ralitrexed for the treatment of advanced colorectal cancer	19
R7	Fludarabine as second line therapy for B-cell lymphocytic leukaemia	18
R11	Trastuzumab for breast cancer	18
R20	Glycoprotein IIb/IIIa antagonists	16
R6	Topotecan for ovarian cancer	16
R14	New drug treatments for rheumatoid arthritis: etanercept and infliximab	16
R12	New drug treatments for juvenile idiopathic arthritis: etanercept	15
R8	Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma	15
R3	Metal-on-metal hip resurfacing arthroplasty for treatment of hip disease	13

- Three reviews used a methodological search filter for the cost-effectiveness searches only.

Searches done in addition to those for clinical and cost-effectiveness studies

Two reviews mentioned that additional database searches were done (apart from those for natural history and background to the condition):

- R2 searched for studies on ease of use and patient/carer preference for compliance with inhaler devices.
- R4 searched for outcome assessment of the adverse effects and safety of nicotine replacement therapy.

Quality assessment of included studies

- All studies assessed the quality of the clinical effectiveness studies in some way. The most common method for assessing RCTs was the Jadad criteria[29] (used in seven reports), but a wide variety of other study designs, and hence quality assessment methods, was used.
- Twelve of the 17 reviews that did identify studies to use for the economic evaluation reported the methods used to assess the quality of those studies. A variety of methods was used, but the most common method was the *BMJ* guidelines for economic appraisals,[30] used in nine TARs.

Analysis of the clinical effectiveness studies

In total, 424 clinical effectiveness studies were included over the 20 reviews. For this study, each was classified into one of three different publication types: published, abstract or unpublished (see Chapter 3).

The results of the classifications were:

- published = 80.0% (339)
- abstract = 11.3% (48)
- unpublished = 8.7% (37).

Table 4 shows the number of studies used for the clinical effectiveness section, broken down by study design and publication type, and sorted in descending order of the total number of studies used.

Number of studies used

Table 4 shows the large range in the total number of studies used in the clinical effectiveness section, ranging from two for etanercept for juvenile idiopathic arthritis (as this was in paediatrics, one would expect very few trials) to 41 for asthma inhaler devices.

Study designs used

It can also be seen that 11 (55%) of the TARs used the 'highest' form of evidence, that is, only from RCTs (which included Phase II RCTs) or systematic reviews. The remaining nine reviews used some data from non-RCT studies.

Table 5 summarises the study designs used across the 20 TARs and 424 included studies. The descriptions used are those given in the TARs. Study designs that were described as a type of RCT (either Phase III, Phase II, cross-over or unspecified) are grouped at the top of the table, and non-RCT designs are in the bottom part of the table. The data show that approximately 72% of the studies were described as a type of RCT and the other 28% were non-RCTs. It is possible that the number of non-RCTs is higher, as some of those studies described as Phase II RCTs may not have been randomised.

Publication types used in TARs

Published studies

Only four reviews (20%) relied exclusively on published data for the clinical effectiveness section. Hence, 80% of TARs used some literature that was not from fully published studies.

Abstracts

It was found that 12 (60%) reviews included at least one abstract. Those reviews that included the highest number were all cancer reviews: R5 (Irinotecan, oxaliplatin and ralitrexed for the treatment of advanced colorectal cancer), which included 14 abstracts (54% of the included studies), and R11 (Trastuzumab for breast cancer), which included ten abstracts (50% of the included studies).

Unpublished studies

Fifty per cent of the reviews included at least one unpublished study. R4 (*Bupropion and nicotine replacement therapy*) had the highest number; it included eight unpublished studies. The next highest was R16 (*Computerised cognitive behaviour therapy*); it included six unpublished articles. R6 (*Topotecan for ovarian cancer*) had three (of a total of seven) studies unpublished.

The 37 unpublished studies were all either RCTs or systematic reviews; 26 (70%) were confidential industry submissions and the other 11 (30%) comprised one dissertation, two posters, five unpublished papers, two NICE appraisals and one interim guidance from NICE.

TABLE 4 *Number of studies used, and study designs included, in the clinical effectiveness section*

Review no.	Short title	Study designs used	Pub.	Abs.	Unpub.	Total
R2	Inhaler devices used in the routine management of chronic asthma in older children	1 systematic review, 19 RCTs, 21 RCT cross-overs	40	1		41
R10	Surgery for people with morbid obesity	28 RCTs, 9 cohort studies with matched controls, 2 systematic reviews	39			39
R20	Glycoprotein IIb/IIIa antagonists	37 RCTs (1 Phase II), 1 systematic review	38			38
R1	Palcitaxel, docetaxel, gemcitabine and vinorelbine in lung cancer	34 RCTs	34			34
R13	Growth hormone in children	11 non-RCTs, 23 RCTs	32	2		34
R3	Metal-on-metal hip resurfacing arthroplasty for treatment of hip disease	1 RCT, 3 systematic reviews, 3 industry submissions (design not specified), 1 guidance for manufacturers and sponsors (design not specified), 17 observational studies, 3 watchful waiting	23		5	28
R17	Home versus hospital or satellite unit haemodialysis for people with end stage renal failure	22 comparative observational studies, 1 cross-over RCT, 4 systematic reviews	26	1		27
R5	Irinotecan, oxaliplatin and ralitrexed for the treatment of advanced colorectal cancer	26 RCTs (4 Phase II, 1 Phase II/III)	11	14	1	26
R11	Trastuzumab for breast cancer	11 RCTs, 8 case series (Phase II), 1 review/case series (Phase II)	9	10	1	20
R18	Ultrasound locating devices for central venous access	20 RCTs	18	2		20
R14	New drug treatments for rheumatoid arthritis: etanercept and infliximab	19 RCTs	10	5	4	19
R4	Bupropion and nicotine replacement therapy for smoking cessation	16 RCTs, 2 systematic reviews	10		8	18
R9	Sibutramine in the management of obesity	16 RCTs	11		5	16
R16	Computerised cognitive behaviour therapy for depression and anxiety	11 RCTs, 2 cohort studies (no comparator group), 1 comparative study, 2 pilot studies	10		6	16
R15	Pegylated liposomal doxorubicin hydrochloride for ovarian cancer	13 Phase II RCTs	4	6	3	13
R19	Routine anti-D prophylaxis for pregnant women who are rhesus-negative	1 community intervention trial, 1 follow-up study, 3 prospective studies, 1 quasi-RCT, 1 RCT, 5 retrospective studies	11	1		12
R7	Fludarabine as second line therapy for B-cell lymphocytic leukaemia	1 RCT, 7 case series	6	2		8
R6	Topotecan for ovarian cancer	7 RCTs	1	3	3	7
R8	Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma	5 prospective case series	5			5
R12	New drug treatments for juvenile idiopathic arthritis: etanercept	1 RCT, 1 extension of RCT	1	1		2

Pub.: published; Abs.: abstract; Unpub.: unpublished.

TABLE 5 *Summary of study designs used*

Study designs (as described in the TARs)	No. of studies	% of studies
RCTs		
RCTs	255	60.1
RCT cross-over	23	5.4
RCT Phase II	18	4.2
RCT Phase III	7	1.7
RCT – withdrawal trial	1	0.2
RCT Phase II/III	1	0.2
Non-RCTs		
Comparative observational study (18 prospective, 3 retrospective, 2 not specified)	23	5.4
Observational study	17	4.0
Case series	15	3.5
Systematic review	14	3.3
Non-RCT	12	2.8
Cohort studies (10 with matched controls; 1 with no comparator group)	11	2.6
Prospective case series	5	1.2
Industry submission (design not specified)	4	0.9
Prospective studies, historical controls	3	0.7
Retrospective study, historical controls	3	0.7
Watchful waiting[a]	3	0.7
Pilot studies (no comparator group)	2	0.5
Community intervention trial	1	0.2
Follow-up study to RCT	1	0.2
Open-label extension of RCT	1	0.2
Quasi-RCT	1	0.2
Retrospective study, geographical controls	1	0.2
Retrospective survey (before and after)	1	0.2
Review/case series – Phase II	1	0.2

[a] Watchful waiting studies refer to a particular variant of prospective observational studies where patients are not treated, but are monitored closely for progression of disease.

Number of participants per review

Table 6 gives the number of participants per TAR. The total number of participants for all 20 TARs was 203,962. There were 184,499 (90.5%) in published studies, 14,453 (7.1%) in abstracts and 5010 (2.5%) in unpublished reports. It should be noted these figures are incomplete as the data on sample size were missing for 19 unpublished RCTs, of which 18 were from company submissions; these 19 RCTs were in six separate reports (R4, R5, R6, R11, R14 and R15).

The proportions of participants in the clinical effectiveness studies were 66% (134,609) from RCTs and 34% (69,353) from non-RCTs. The latter includes a wide range of different study designs (as shown in *Table 5*).

Table 6 shows a very wide range in the number of participants per review, from 69 for R12 (*New drug treatments for juvenile idiopathic arthritis: etanercept*) to 97,570 for R20 (*Glycoprotein IIb/IIIa antagonists*). Both reviews only included RCTs or systematic reviews. The differences between the size of the evidence base for various reviews will be determined by a number of factors, such as the scope of the review, how new the intervention is, the frequency of the disease, the relative amount of research funds for each disease and the amount of unpublished trial data that manufacturers will make available. The size of the effect can also be important, as it takes larger samples to prove a small effect.

The median number of participants per TAR was found to be 2787, the mean was 10,198 and the IQR was 865 to 6813.

Age of clinical effectiveness studies used

The ages of the studies used were analysed to provide some insight into how far back one needs to search. The median age of studies used in the clinical effectiveness section of the TARS was 2 years, the mean was 4.5 years and the IQR was 1 to 6 years.

TABLE 6 *Number of participants per review*

Review no.	Short title	No. of participants
R20	Glycoprotein IIb/IIIa antagonists	97,570
R19	Routine anti-D prophylaxis for pregnant women who are rhesus-negative	44,833
R17	Home versus hospital or satellite unit haemodialysis for people with end stage renal failure	13,478
R1	Palcitaxel, docetaxel, gemcitabine and vinorelbine in lung cancer	7,457
R10	Surgery for people with morbid obesity	7,031
R5	Irinotecan, oxaliplatin and ralitrexed for the treatment of advanced colorectal cancer	6,160
R2	Inhaler devices used in the routine management of chronic asthma in older children	5,647
R9	Sibutramine in the management of obesity	4,162
R3	Metal-on-metal hip resurfacing arthroplasty for treatment of hip disease	3,085
R4	Bupropion and nicotine replacement therapy for smoking cessation	2,925
R13	Growth hormone in children	2,648
R14	New drug treatments for rheumatoid arthritis: etanercept and infliximab	2,340
R18	Ultrasound locating devices for central venous access	1,836
R7	Fludarabine as second line therapy for B-cell lymphocytic leukaemia	1,393
R16	Computerised cognitive behaviour therapy for depression and anxiety	908
R11	Trastuzumab for breast cancer	850
R6	Topotecan for ovarian cancer	709
R15	Pegylated liposomal doxorubicin hydrochloride for ovarian cancer	474
R8	Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma	387
R12	New drug treatments for juvenile idiopathic arthritis: etanercept	69

The study with the lowest mean age of studies used, 0.8 years, was R8 (*Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma*); the highest was R19 (*Routine anti-D prophylaxis for pregnant women who are rhesus-negative*), with a mean age of 12.3 years. This reflects the ranges in the ages of the technology being assessed.

Language of studies used in the TARs

All studies, except for one in French, were published in English.

Proportion of trials indexed in the major databases

The proportion of the studies that could be found indexed in the six databases: CCTR (C), MEDLINE (M), EMBASE (E), SCI (S), BIOSIS (B) and ASCO (A), both individually and cumulatively, was measured.

Table 7 shows the proportion of studies, using all publication types as the denominator, that would have been found if CCTR, MEDLINE and EMBASE had each been searched individually.

TABLE 7 *Numbers of all publication types found in the databases CCTR (C), MEDLINE (M) and EMBASE (E)*

Database	No. found (of all publication types)	% (of 424 studies)
M	320	75.5
E	304	71.7
C	194	45.8

It can be seen that MEDLINE had the highest percentage of studies (75.5%), followed by EMBASE (71.7%). The reason for CCTR being much lower (45.8%) is that this database only covers CCTs, and as shown earlier only about 72% of studies in this sample are RCTs.

Percentage of published studies indexed in CCTR, MEDLINE and EMBASE

It was found (see above) that 8.7% (37/424) of all clinical effectiveness studies were unpublished, and therefore by definition would not be found indexed in any of these databases. It was therefore

TABLE 8 *Published studies found by searching combinations of CCTR (C), MEDLINE (M) and EMBASE (E)*

Database	No. found (published studies only)	% (of 339 studies)
Single database		
M	320	94.4
E	304	89.7
C	189	55.8
Two databases		
M + E	333	98.2
C + M	326	96.2
C + E	316	93.2
Three databases		
M + E + C	333	98.2

TABLE 9 *Number of studies (published or abstracts) found in searches of six databases*

Databases	No. retrieved (published or abstracts)	% (of total 387 studies)
Single database		
M	320	82.7
E	304	78.6
C	194	50.1
Two databases		
M + E	333	86.1
M + C	331	85.5
C + E	321	83.0
Three databases		
M + E + C	338	87.3
Four databases		
M + E + C + S	356	92.0
Five databases		
M + E + C + S + B	360	93.0
Six databases		
M + E + C + S + B + A	380	98.2

M: MEDLINE; E: EMBASE; C: CCTR; S: SCI; B: BIOSIS; A: ASCO.

decided to determine the proportion of the studies published in full that would be found in the three databases, as this would give a more realistic measure of their coverage. *Table 8* includes data for the three databases, both individually and cumulatively, in all combinations.

Table 8 shows that MEDLINE still gave the highest retrieval (94.4% of published studies), with EMBASE slightly less (89.7%), followed by CCTR (55.8%). The highest yield for two databases is the combination of MEDLINE and EMBASE, which together retrieved 98.2% of included studies; adding EMBASE to the MEDLINE search retrieved an extra 13 (3.8%) studies. Adding CCTR to the search of MEDLINE and EMBASE did not retrieve any additional studies.

Percentage of published studies or abstracts indexed in six databases

Finally, the percentage of studies that were classified as either published or abstract was calculated. Nearly all of the meeting abstracts found were published in journal supplements. Therefore, it would be useful to examine the collective proportion of these two publication types found in the electronic databases.

As outlined in the Methods section (Chapter 3), all studies classified as published or abstracts were checked to see whether they were indexed in CCTR, MEDLINE and EMBASE. Any studies not found in any of these three databases were then checked to see whether they were indexed in SCI or BIOSIS. Any studies still not found in any database were checked to see whether they were found in the database of ASCO abstracts.

Table 9 shows the data obtained from these searches, for individual databases, and the cumulative numbers from progressively searching additional databases.

The histogram in *Figure 1* displays the percentages found individually in the three databases. As the numbers found in MEDLINE and EMBASE (320 and 304, respectively) are unchanged from *Tables 8* and *9*, it can be seen that no abstracts were found in these databases. However, CCTR has 194 in *Table 9* (compared with 189 in *Table 8*), which means that an extra five abstracts were found.

Figure 2 shows the cumulative percentages of published articles or abstracts retrieved from six databases.

Analysis of the extra studies retrieved when searching beyond MEDLINE

Figure 2, which is based on the data in *Table 9*, shows the cumulative number of studies found when the databases were searched in the order MEDLINE, EMBASE, CCTR, SCI, BIOSIS and ASCO. Many other orders of searching are possible, but this order was chosen as it gave the highest yield of studies at each step. (Note that the ASCO search only applies to cancer TARs.)

It is noted that the extra numbers retrieved at each step are small and that there is a large overlap of confidence intervals for each subsequent step, indicating that a different sample of reports could yield a different order for these databases.

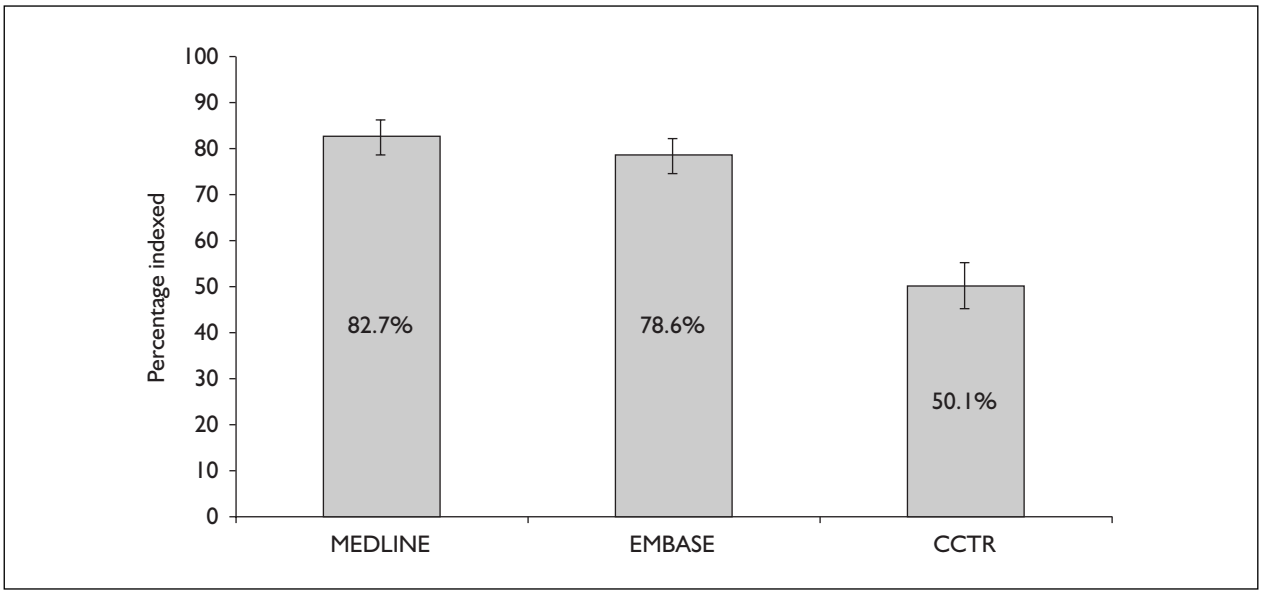


FIGURE 1 *Proportion of articles (published or abstracts) indexed in MEDLINE, EMBASE and CCTR. Error bars represent the 95% confidence intervals (CI).*

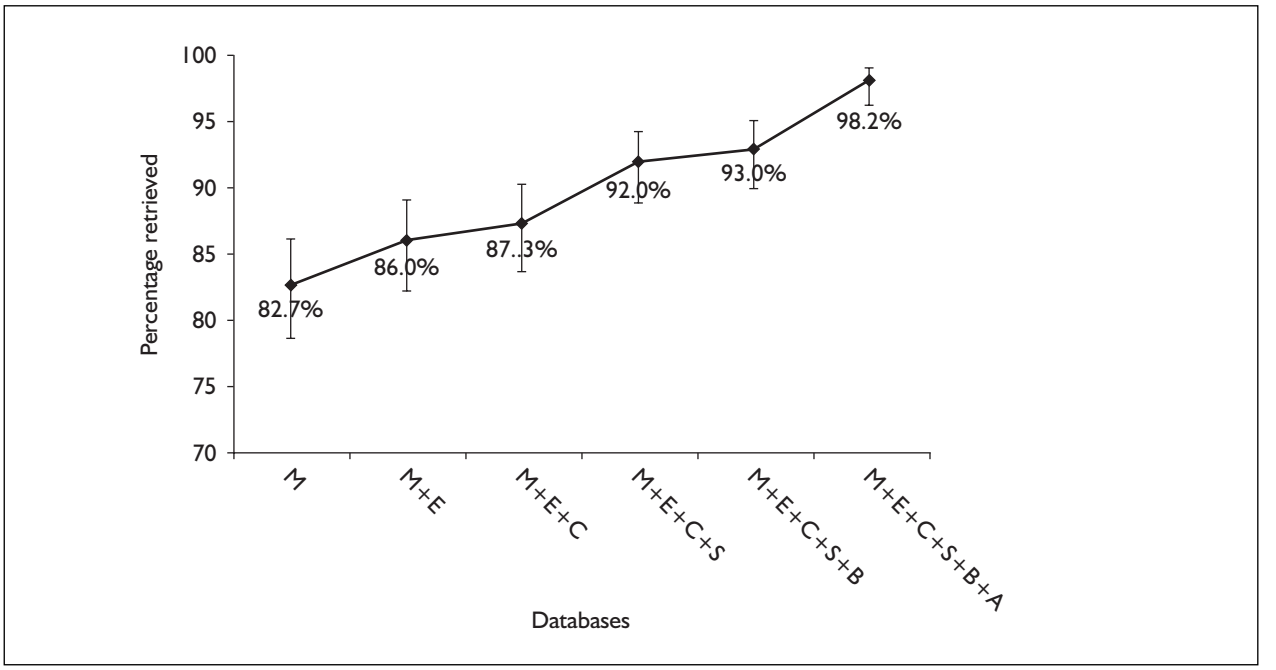


FIGURE 2 *Cumulative percentages of published articles or abstracts retrieved from six databases. Error bars represent the 95% CI.*

The following section gives details of the extra studies retrieved at each step following Search 1 – the MEDLINE search.

Search 2. Adding EMBASE to the MEDLINE search

An extra 13 studies (spread over seven reviews) were retrieved. They came from ten different journals; nine were RCTs and four were observational studies.

A check was done to determine whether any of the ten journals are now indexed in MEDLINE. It was found that:

- three of the journals are now indexed by MEDLINE: *International Journal of Oncology* (from 1998), *Obesity Surgery* (from 1997) and *Oncologist* (from 1999)
- the other seven are still unique to EMBASE: *Behaviour and Cognitive Psychotherapy, Journal of*

TABLE 10 *Number of participants per study: EMBASE versus all studies*

	Median	Mean	IQR	Min.	Max.
All studies (n = 324)	110	630	45 to 312	9	16,588
EMBASE only (n = 12; 1 missing)	6	134	2 to 6	0	16

TABLE 11 *Age of studies (in years) found in EMBASE versus all studies*

	Median	Mean	IQR	Min.	Max.
All studies (n = 423)	2.0	4.5	1 to 6	0	24
EMBASE only (n = 13)	5.0	5.5	2 to 6	0	16

Mental Health UK, Dialysis and Transplantation, Pediatric Asthma Allergy International, Journal of Pharmaceutical Medicine, Current Therapeutic Research – Clinical and Experimental, and Hip International.

Only two of the seven journals still unique to EMBASE are exclusively pharmaceutical journals.

Characteristics of studies found in EMBASE but not MEDLINE

The number of participants per study (for which data were available) and the age of studies found in EMBASE but not MEDLINE were compared with the rest of the studies (for which data were available). The results are shown in *Tables 10* and *11* respectively. There was a tendency for studies unique to EMBASE to be older and smaller, but the number of such studies was too small to allow any meaningful comparison.

Search 3. Adding CCTR to the MEDLINE + EMBASE search

An extra five studies were found in CCTR. These were used in three different reviews. All were abstracts, and all RCTs.

- One was an *ASCO* abstract, and in the ASCO database.
- One was in a supplement to *Thorax*, and not indexed in BIOSIS or SCI (it had been identified by Cochrane handsearchers).
- Three were in a supplement to the *European Journal of Cancer* (and were also indexed in BIOSIS and SCI). However, the full abstracts of all three were in CCTR, whereas SCI and BIOSIS did not include the abstract itself, only the bibliographical details.

Search 4. Adding SCI to the MEDLINE + EMBASE + CCTR search

An extra 18 studies (14 of which were RCTs) were identified in the SCI search; two were published, 16 were abstracts. They appeared in seven separate reviews.

- The two articles published in full were in separate journals; one in *Endocrinologist*, one in *European Heart Journal Supplement*.
- The 16 abstracts were in the supplements of the following journals: six in *European Journal of Cancer*, three in *Arthritis and Rheumatism*, two in the *Annals of Rheumatic Diseases*, and one each in *Annals of Oncology, Blood, British Journal of Cancer, British Journal of Haematology and Pediatric Research*. (Hence, nine of 16 abstracts were in oncology journals.)

Thus, it would appear that there are some fully published articles (plus a large number of abstracts) that are in SCI, but not in MEDLINE or EMBASE.

Search 5. Adding BIOSIS to the MEDLINE + EMBASE + CCTR + SCI search

An extra four studies, from four separate reviews, were found in BIOSIS. All were abstracts, of which two were RCTs.

- The abstracts came from four different journals: *Cancer Investigation, Hormone Research, Nephrology Dialysis Transplantation* and *British Journal of Obstetrics* and *Gynaecology*.

A calculation was done to determine the proportion of the 18 articles found above in SCI that were also found in BIOSIS. The results

TABLE 12 *Articles still not retrieved after searching the six databases*

Review no.	Journal	Format	Year	Study design
R10	*CRD Report 10*	Published	1997	Systematic review
R12	*EULAR Abstract*	Abstract	2000	Open-label extension of RCT
R16	*Depression*	Published	1993	RCT
R17	*Perit Dial Bull*	Published	1988	Comparative observational study
	Loss Grief and Care	Published	1991	Comparative observational study
R18	*Crit Care Med*	Abstract	1996	RCT
	Acad Emerg Med	Abstract	2001	RCT

showed that ten of 18 (55.6%) of articles were in both databases. This gives an overlap between SCI and BIOSIS of 45% for the extra 22 articles identified after searching MEDLINE + EMBASE + CCTR.

Search 6. Adding ASCO to the MEDLINE + EMBASE + CCTR + SCI + BIOSIS search

An extra 20 ASCO abstracts were found in the ASCO database. These studies were included in four separate reviews of cancer drugs: R5, R6, R11 and R15. (ASCO is, by definition, only relevant to cancer reviews.)

Articles still not retrieved after searching the six databases

Seven articles (three abstracts and four published) were not found in the six databases. They are summarised in *Table 12*. They came from five different reviews; three were RCTs.

The sources in which these other studies could be found were checked. It was found that *CRD Report 10* was indexed in the HTA database on the Cochrane Library, and the article in *Loss Grief and Care* was found indexed in PsycINFO. The article in *Depression* was cited by another article used in the clinical effectiveness section. None of the other four articles was found indexed anywhere, despite extensive searching of a number of databases (see *Chapter 3*). It is assumed that they were probably identified through reference lists or from the industry submissions.

Studies in MEDLINE, but not in EMBASE

An investigation was done of the studies that were found in MEDLINE but not in EMBASE. There were 29 articles, which came from 20 different journals; seven were pre-1980 studies and came from five different journals. The version of EMBASE used in this study for checking only went back to 1980, so one would not expect to find these seven articles indexed. Below is a breakdown of the sources of the 22 post-1980 articles unique to MEDLINE.

- Five of the journals are now indexed by EMBASE: *MD Computing* (from 2000), *Pediatric Pulmonology* (from 1996), *Pediatric Transplantation* (from 1999), *Transfusion Medicine* (from 2001) and *Transfusion Medicine Reviews* (from 1996).
- Seven are still not indexed: *Clinical Nursing Research, International Journal of Pediatric Nephrology, Journal of the American Association of Nephrology Nurses and Technicians, Nephrology Nurse, Obesity Research, Transactions – American Society for Artificial Internal Organs* and *Cochrane Reviews*.
- One was in *Acta Chirurgica Scandinavica, Supplement*, which is no longer indexed in EMBASE.
- There was one from each of three journals, *Annals of Oncology, Journal of Clinical Endocrinology and Metabolic Disorders* and *Journal of Anxiety Disorders* which are all indexed by EMBASE, but for some reason these particular articles were missing.

Overlap between MEDLINE and EMBASE for the clinical effectiveness studies

The number of records that are found in both MEDLINE and EMBASE (the overlap) was calculated. It was found (see *Table 9*) that a combined total of 333 unique articles was retrieved when searching MEDLINE and then EMBASE. It was also found that 291 of these articles were common to both databases; therefore, the overlap for studies included in the clinical effectiveness section was calculated as 87.4% (95% CI 83.4 to 90.5%).

Delay in inclusion of studies in CCTR

It was found (see *Table 9*) that 194 of the total 424 articles were found in CCTR when the version of the Cochrane Library used for the TAR was searched. A second search was done of the

Cochrane Library 2002, Issue 4 (the current issue) for all articles that were not in the first CCTR search. It was found that 44 of these articles had now been added to the later version of CCTR.

The 'ages' of these 44 studies were determined. This is the difference in the number of years between the year of publication of the version of the Cochrane Library searched for the TAR, and the year in which the study was published. The results for the number of years' delay before appearing in CCTR were: 34 were 0 years, nine were 1 year and one was 2 years (published in the *International Journal of Obesity and Related Metabolic Disorders*).

It is recognised that this is most likely be an underestimate of the delay, as more studies could appear in subsequent issues of the Cochrane Library. However, it does provide some insight into the delay between studies being published and being included in CCTR.

Analysis of RCTs not in CCTR

This section analyses published studies that were described in the TAR as being an RCT (of any sort), but were not indexed in the Cochrane Library 2002, Issue 4. Only 15 articles fitted these criteria. The aims were to determine: (1) whether they were RCTs that would fit the CCTR criteria for inclusion and, if so, (2) possible reasons why they may not have been in CCTR.

The full papers of these 15 articles were obtained and examined to determine whether they were really RCTs. The results were as follows.

- Six of the 15 turned out not to be primary reports of RCTs; three of the Phase II RCTs were not RCTs as they had no control group, and three of the others were summaries of RCTs presented at meetings, i.e. they were secondary reports that did not report new data.
- Nine appeared to be 'true' RCTs. They came from nine different journals; one was published in 1980 and one in 1978; the other seven were all published since 1993.

Characteristics of the nine RCTs and possible reasons why they were not in CCTR

- Five were in MEDLINE (which is searched by the Cochrane Collaboration as a source of trials using a three-phase search filter to retrieve RCTs, with phase I being a highly specific strategy, and the other two phases progressively more sensitive[31]). Two of the five would have been found using phase I of this strategy and one using phase II, but the full papers would need to be read to confirm that they were RCTs (as the abstracts did not allow one to determine whether the study was an RCT). The other two would not have been picked up by any phase of the strategy, and one could only know they were RCTs from reading the methods section of the full paper.
- One (published in 1999) was in EMBASE (which is also searched by the Cochrane Collaboration as a source of trials), and had the phrase 'randomized, multicenter, double-blind, parallel-group study' in the abstract, so it would be found with a highly specific search filter for RCTs.
- Two were indexed in SCI only; they had the word 'randomly' or 'randomised' in the abstract, so both would be retrieved with a highly specific search filter for RCTs.
- One was not indexed in any database. It appeared in *Depression*, Volume 1, 1993, and was not yet indexed as it was a new journal. This journal is continued, called *Depression and Anxiety*, and is now indexed in MEDLINE.

Therefore, of the nine RCTs not in CCTR, five would have been retrieved using a highly specific search filter for RCTs (i.e. in these cases, using the search *random**) run against MEDLINE, EMBASE and SCI, two would have been retrieved by phase II of the strategy, two would not have been retrieved by any phase, and one was not indexed in any database checked in this study.

Search terms for finding non-RCTs

In total, 118 non-RCTs were used in this study (101 published, ten abstracts and seven unpublished). Twelve of these were systematic reviews, which left 89 non-RCT studies. An analysis was done of the indexing terms of each of these articles, by examining the bibliographic records downloaded from the database in which they were indexed.

A combination of the keyword terms (terms in the Title, Abstract, or Subject Heading fields) combined with the MEDLINE Publication Type (PT) terms; that is,

> case control studies OR clinical trial* OR cohort OR comparative OR comparison OR continuing study OR control* OR cross sectional studies OR drug evaluation OR epidemiology OR follow-up studies OR intervention study OR longitudinal studies OR long-term OR matched OR phase II OR pilot projects OR prospective studies OR retrospective studies OR survival OR treatment outcome OR clinical trial (PT) OR comparative study (PT) OR multicenter study (PT)

TABLE 13 *Search terms used to describe non-RCTs in bibliographic records*

Study design	No.	Search terms	No. found
Case series	14	clinical trial (PT) OR multicenter study (PT) OR clinical trial* OR drug evaluation	14
Cohort study with matched controls	9	control*	9
Comparative observational study	22	compar* OR comparative study (PT)	10
Non-RCT	5	controlled clinical trial (PT) OR control*	5
Non-RCT – retrospective	3	cohort OR control* OR retrospective studies	3
Observational study	17	long-term OR follow-up OR prospective OR treatment outcome OR compare*	14
Retrospective study, historic controls	3	clinical trial (PT) AND followed	2
Watchful waiting	3	prospective stud* OR cross-sectional studies	3
Total	76		60

* Represents a truncation symbol.

retrieved 76 of 89 (85%) of the non-RCT studies found; hence, there were still 13 articles that did not have any of these terms in them, or any other terms that seemed to be related to study design. It would appear that these 13 articles would only be identified by a very sensitive search on the intervention, without a methodological filter, and then by scanning all the records for those that might look relevant. This would be a time-consuming exercise.

Table 13 gives examples of the terms used for some of the different study designs, and their retrieval rate (only those studies with three or more of a particular study design are shown). This covers 76 of the 89 studies (although all 89 were analysed). All search terms are keyword terms, unless indicated as being from the Publication Type field, that is, with PT in brackets after the search term. (The results of such a search of MEDLINE could vary slightly with different search interfaces.)

Analysis of published studies not in MEDLINE

It was found (see *Table 8*) that 19 of the 339 (5.6%) of the published studies were not in MEDLINE. These were spread over nine TARs and came from 16 separate sources; 13 of 19 were in EMBASE (described in detail in the previous section). Of the 19 studies, 12 were RCTs and seven were other study designs. *Table 14* gives details of the non-MEDLINE articles, including their quality score as given in the review (where available), the number of participants in the study (sample size) and the percentage of the total number of participants that were included.

The non-MEDLINE article that contributed the highest proportion of the sample size (8.4%) was in the review R2 (Inhaler devices used in the routine management of chronic asthma in older children), published in *Current Therapeutic Research – Clinical and Experimental* (which was indexed in EMBASE). The article also appeared to be of a reasonable quality (Jadad score of 3), which compared favourably with the other trials in this review that were given Jadad scores (two had a Jadad score of 4, 14 had a Jadad score of 3, five had a Jadad score of 2 and six had a Jadad score of 1). *Current Therapeutic Research* is also indexed in SCI, and has an impact factor of 0.562.

One problem with using the Jadad score for comparing trials is that with interventions where blinding is impossible (e.g. gastric surgery for obesity, different methods of dialysis or different inhalers) the maximum Jadad score will be 3. In these cases, the score can be reported as 3/3. In some cases, the lack of blinding does not matter, because the results are in terms of hard end-points, the measurement of which is not susceptible to bias.

The next most significant studies in terms of study size were non-RCTs, from *Journal of Mental Health UK* and *Dialysis and Transplantation*, which contributed 3.6% and 3.4% of the total sample size, respectively; both articles were indexed in EMBASE. The remaining 13 articles (for which an equivalent calculation was possible) all contributed less than 2.5% or less of the total sample size. Nine of the 13 were indexed in EMBASE; three (from *Loss Grief and Care*, *Peritoneal Dialysis Bulletin* and *Depression*) were not indexed in

TABLE 14 *Characteristics of published studies not indexed in MEDLINE*

Review no.	Journal	Quality score	Design	Sample size	% of total no. in review
R1	*Int J Oncol*	Jadad 1/5	RCT	52	0.7
R1	*Int J Oncol*	Jadad 2/5	RCT	69	0.9
R2	*J Pharmac Med*	Cochrane A	RCT cross-over	25	0.4
R2	*Pediatr Asthma Allergy Immunol*	Jadad 1	RCT cross-over	13	0.2
R2	*Curr Ther Res Clin Exp*	Jadad 3	RCT	473	8.4
R3	*Hip Int*	n.a.	Observational study	51	1.7
R10	*Obes Surg*	n.a.	RCT	42	0.6
R10	*Obes Surg*	n.a.	RCT	106	1.5
R10	*CRD Report 10*	n.a.	Systematic review		
R11	*Oncologist*	n.a.	RCT	469[a]	
R13	*Endocrinologist (Suppl)*	Jadad 2/5	RCT	54	2.0
R16	*Depression*	Jadad 2	RCT	22	2.4
R16	*Behaviour and Cognitive Psychotherapy*	Jadad 2	RCT	23	2.5
R16	*Journal of Mental Health UK*	n.a.	Pilot study (no comparator group)	33	3.6
R17	*Dialysis and Transplantation*	17/27	Comparative observational study – prospective	454	3.4
R17	*Peritoneal Dialysis Bulletin*	9/27	Comparative observational study – prospective	194	1.4
R17	*Loss Grief and Care*	13/27	Comparative observational study – prospective	42	0.3
R17	*Dialysis and Transplantation*	4/27	Comparative observational study – prospective	269	2.0
R20	*Eur Heart J Suppl*		RCT Phase III	10,948[b]	

[a] One of nine articles of a trial with 469 patients.
[b] One of four articles of a larger trial with 10,948 patients.
n.a.: not available.

TABLE 15 *Quality characteristics of RCTs not in MEDLINE*

Review no.	Journal	Year	Design	Size	Allocation concealment	Blinding
R1	*Int J Oncol*	1995	RCT	69	Not reported	Not reported
R1	*Int J Oncol*	1997	RCT	52	Not reported	Not reported
R2	*Curr Ther Res Clin Exp*	1999	RCT	473	Not reported	Double-blind
R2	*Pediatr Asthma Allergy Immunol*	1995	RCT cross-over	13	Not reported	Not possible
R2	*J Pharmac Med*	1995	RCT cross-over	25	Not reported	Double-blind
R10	*Obes Surg*	1995	RCT	106	Not reported	Not possible
R10	*Obes Surg*	1995	RCT	42	Not reported	Not possible
R11	*Oncologist*	1998	RCT		Not reported	Not reported
R13	*Endocrinologist (Suppl)*	2000	RCT	54	Not reported	Not reported
R16	*Behaviour and Cognitive Psychotherapy*	2001	RCT	23	Not reported	Not feasible
R16	*Depression*	1993	RCT	22	Not reported	Outcome assessors blinded
R20	*Eur Heart J Suppl*	1999	RCT phase III	10,948	Not reported[a]	Double-blind[a]

[a] One of four studies of a large trial (with 10,948 patients); methods were reported elsewhere.

any database checked, and one [from *Endocrinologist (Supplement)*] was indexed in SCI.

An analysis of the full papers of all 12 RCTs in *Table 15* was done to determine their allocation concealment and blinding. Both of these characteristics have been shown to be strongly associated with trial quality; trials with inadequate or unclear concealment of allocation, and trials that are not double-blind, overestimate treatment effects by about 30% and 15%, respectively.[32]

The results show that none of the 12 RCTs reported on allocation concealment and only four reported that the trial was double-blind; however, in some cases blinding was not feasible (one trial comparing two different asthma inhalers, and two trials on gastric surgery). However, as it was beyond the scope of this study to conduct a comparative analysis of all trials, it is not known whether these trials are of poorer quality than the other included trials that were indexed in MEDLINE.

Impact factors and size of study

There was no correlation between impact factors and number of patients within each of the 20 reviews (data not shown).

Analysis of the cost-effectiveness studies

Two types of search appeared to be done in the economics sections:

- literature searches for existing economic evaluations
- searches for data to support a new economic model, usually done because an adequate one does not exist already. (One problem may be that existing models may be from outside the UK and may not be immediately applicable to a UK context.)

A total of 251 studies was used in the economics sections. These were subdivided into:

- 130 used for the economic evaluations
- 121 used in the economic modelling.

Eight of 130 articles used in the economic evaluation section were also used in the clinical effectiveness section. The cost-effectiveness studies were classified into five different publication types: published, abstract, grey literature, reference source and unpublished.

Table 16 shows the number of studies used per review for both the economic evaluation and economic modelling sections, sorted in descending order for the number used in the economic evaluations. Only 19 TARs were included in this section, as one, R11 (Trastuzumab for breast cancer), did not attempt to determine cost-effectiveness.

It can be seen that R20 included the most articles (26) in the economic evaluation section, and three reviews (R3, R9 and R12) only cited one study each. Two reviews (R13 and R18) that searched for economic evaluations did not find any studies.

The median number of studies used in the economic evaluation was 4.0, the mean was 7.7 and the IQR was 2 to 13.

Fourteen reviews included an economic model. The review that cited the most studies (19) in the economic modelling was R12 (New drug treatments for juvenile idiopathic arthritis: etanercept) and the one that cited the least was R2 (Inhaler devices used in the routine management of chronic asthma in older children), which used two studies.

The median number of articles cited for the economic modelling was 9.0, the mean was 8.6 and the IQR was 2 to 19.

Analysis of articles used in the economic evaluations

The publication types of the 130 articles were:

- published = 95 (73.1%)
- unpublished = 31 (23.8%)
- abstracts = 2 (1.5%)
- grey literature = 2 (1.5%).

Table 17 shows that MEDLINE and EMBASE are equally effective in retrieving studies for the economic evaluation (64.6% each), and that less than half of that proportion (30.8%) were indexed in NHS EED.

As 31 articles were unpublished (all apart from one were from company submissions) and two were from grey literature, one would not expect them to be found in any of these databases. Therefore, only those articles that were classified as either published in full, or as abstracts, were used to calculate the proportion that could be found in the five databases (MEDLINE, EMBASE, NHS EED, SCI and ASCO). The data are shown in *Table 18*. The percentages retrieved individually

TABLE 16 *Number of studies used per review in the economic evaluation and modelling*

Review no.	Short title	Economic Evaluation	Modelling
R20	Glycoprotein IIb/IIIa antagonists	26	n.d.
R17	Home versus hospital or satellite unit haemodialysis for people with end stage renal failure	19	5
R1	Palcitaxel, docetaxel, gemcitabine and vinorelbine in lung cancer	16	10
R4	Bupropion and nicotine replacement therapy for smoking cessation	16	4
R5	Irinotecan, oxaliplatin and ralitrexed for the treatment of advanced colorectal cancer	10	2
R19	Routine anti-D prophylaxis for pregnant women who are rhesus-negative	9	6
R2	Inhaler devices used in the routine management of chronic asthma in older children	8	2
R16	Computerised cognitive behaviour therapy for depression and anxiety	5	8
R6	Topotecan for ovarian cancer	4	n.d.
R10	Surgery for people with morbid obesity	4	14
R15	Pegylated liposomal doxorubicin hydrochloride for ovarian cancer	3	6
R14	New drug treatments for rheumatoid arthritis: etanercept and infliximab	3	14
R7	Fludarabine as second line therapy for B-cell lymphocytic leukaemia	2	n.d.
R8	Rituximab as third-line treatment for refractory or recurrent Stage III or IV follicular non-Hodgkin's lymphoma	2	n.d.
R3	Metal-on-metal hip resurfacing arthroplasty for treatment of hip disease	1	11
R9	Sibutramine in the management of obesity	1	n.d.
R12	New drug treatments for juvenile idiopathic arthritis: etanercept	1	19
R18	Ultrasound locating devices for central venous access	0	10
R13	Growth hormone in children	0	10

n.d., not done.

TABLE 17 *Percentage of economic evaluations (all publication types) found in three databases*

	No.	% in database
MEDLINE	84	64.6
EMBASE	84	64.6
NHS EED	40	30.8

in MEDLINE, EMBASE and NHS EED are also shown in *Figure 3*.

Overlap between MEDLINE and EMBASE for the economic evaluation studies

The number of records that are found in both MEDLINE and EMBASE (the overlap) was calculated. It was found (see *Table 18*) that there was a combined total of 90 unique articles retrieved when searching MEDLINE and then EMBASE. A search of MEDLINE and EMBASE retrieved 78 articles common to both; therefore, the overlap for studies included in the economic

TABLE 18 *Number of economic evaluations (published or abstracts) found from cumulative database searches of five databases*

	No.	% in database (of 97 total)
Single database		
M	84	86.6
E	84	86.6
N	39	40.2
Two databases		
M + E	90	92.8
N + E	88	90.7
M + N	87	89.7
Three databases		
M + E + N	92	94.8
Four databases		
M + E + N + S	93	95.9
M + E + N + A	94	96.9
Five databases		
M + E + N + S + A	95	97.9

M, MEDLINE; E, EMBASE; N, NHS EED; S, SCI; A, ASCO.

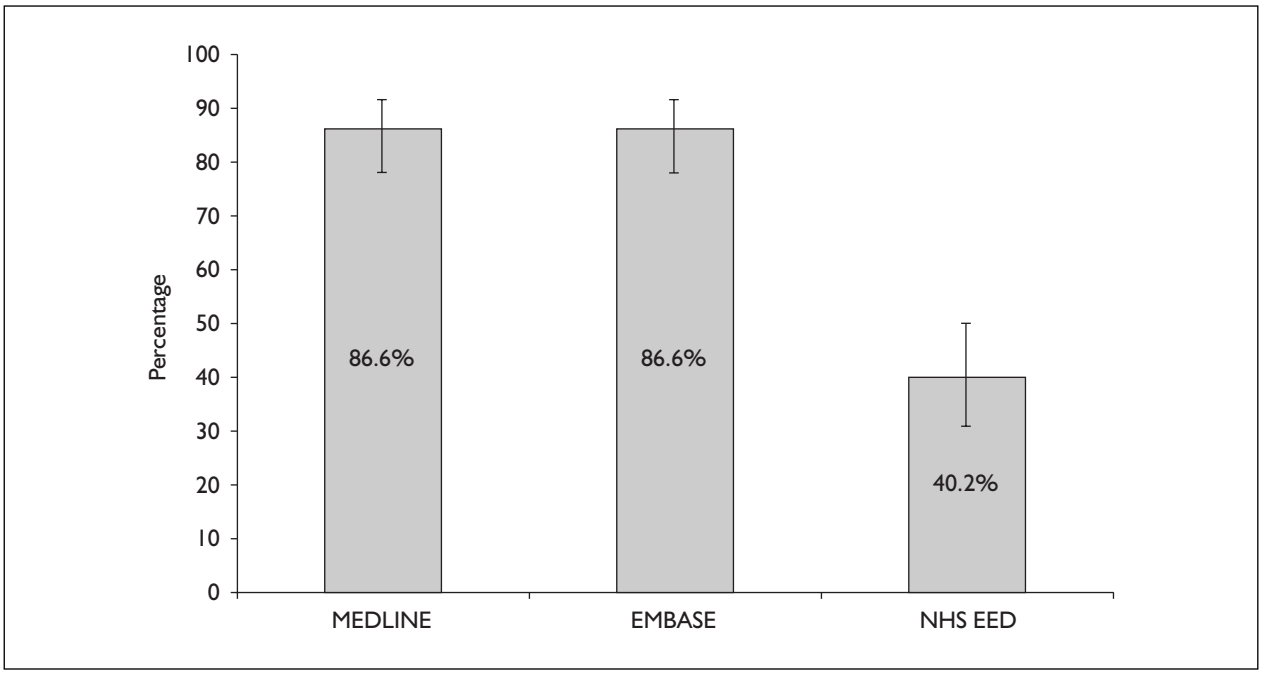


FIGURE 3 *Percentages of economic evaluations (published or abstracts) in three databases. Error bars represent the 95% CI.*

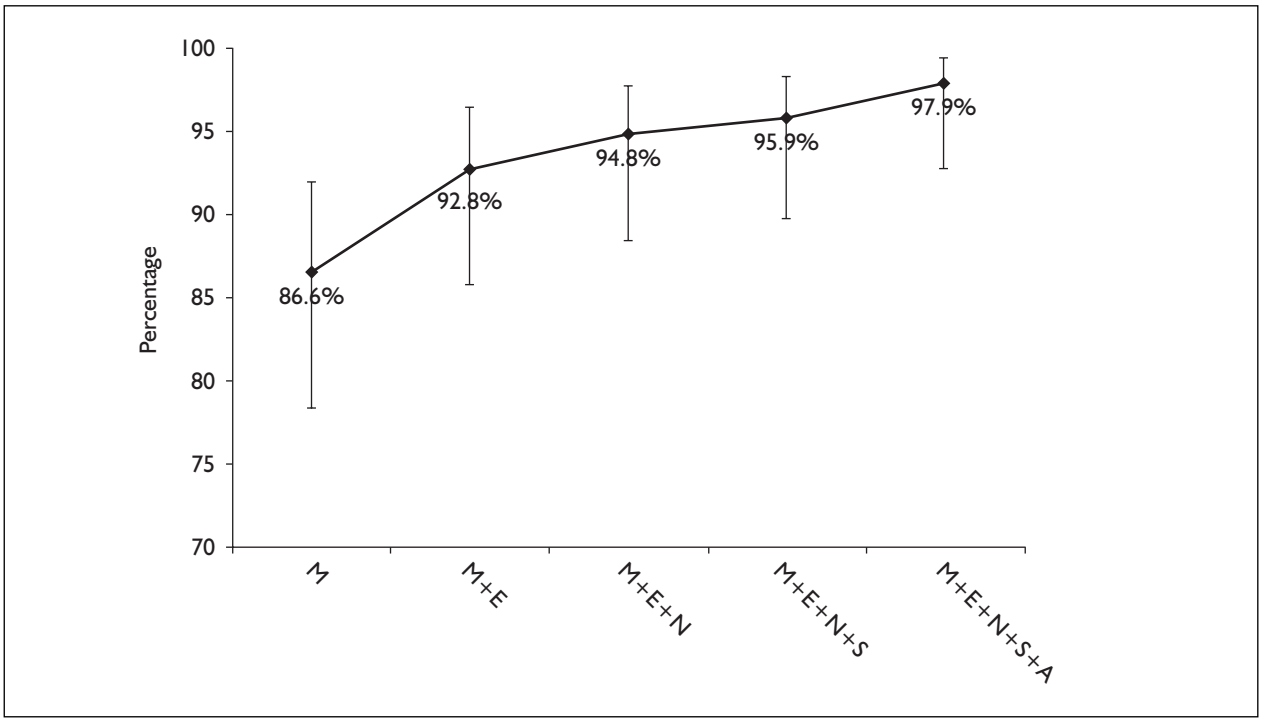


FIGURE 4 *Cumulative proportions of economic evaluations (published or abstracts) retrieved. Error bars represent the 95% CI.*

evaluation sections was calculated as 86.7% (95% CI: 78.1 to 92.2%).

The cumulative numbers retrieved are shown in the graph in *Figure 4*, using the order shown of MEDLINE, EMBASE, NHS EED, SCI and ASCO. On the basis of the data shown in *Table 18*, this order of searching would appear the most effective. Once again, it is noted that the extra numbers retrieved at each step are small and that there is a large overlap of confidence intervals for each subsequent step, indicating that a different sample of reports could yield a different order for these databases.

Analysis of the extra economic evaluation studies retrieved when searching beyond MEDLINE

The following section gives a breakdown of the additional studies retrieved at each step following an initial search of MEDLINE, which will be called Search 1.

Search 2. Adding EMBASE to the MEDLINE search

An extra six studies were found in EMBASE. They came from five different journals and were in five separate reviews:

- three from *British Journal of Medical Economics* (which is now renamed *Journal of Medical Economics*)
- one each from *Journal of Medical Economics, Obesity Surgery* (now indexed in MEDLINE) and *Seminars in Dialysis*.

Search 3. Adding NHS EED to the MEDLINE + EMBASE search

An extra two studies were found in NHS EED:

- one from *Applied Economics* (which was indexed in SSCI)
- one from *European Heart Journal Supplement* (indexed in SCI).

Step 4. Adding SCI to the MEDLINE + EMBASE + NHS EED search

- One extra study, from *Seminars in Dialysis*, was found in SCI.

Step 5. Adding ASCO to the MEDLINE + EMBASE + NHS EED + SCI search

- An additional two ASCO Abstracts were found from the ASCO database.

After searching these five databases, there were two studies that had still not been found. One was from *Home Hemodialysis International* and the other from the journal *Value in Health*. It was likely that the latter article was cited by at least one of the clinical effectiveness articles, as there was an author in common.

The 'added value' of NHS EED

Only two studies were not found in MEDLINE, EMBASE or SCI. One was published in *Applied Economics* (which is indexed in SSCI) and one in the grey literature, a *HERU Discussion Paper*.

Proportion of economic evaluation studies identified using the terms cost* or economic*

An analysis was done to determine what proportion of the economic evaluation studies would be retrieved using some variant of the words 'cost' or 'economic', to determine the effectiveness of a very simple methodological filter for economic evaluation studies.

Four of the 95 published studies were also used in the clinical effectiveness section, so presumably had been identified by the clinical effectiveness search. This left 91 published economic evaluation studies for analysis.

The search terms found in the three fields: Title (TI), Abstract (AB) and Subject Heading (SH), and the numbers that would be retrieved from each of the three fields if searched individually, are summarised below:

- cost* OR economic* in TI = 84
- cost* OR economic* in AB = 8
- cost* OR economic* in SH = 4.

A keyword search (which searches across all three fields) of: cost* OR economic* would retrieve 88 of a total of 91 articles. Note that some articles have the terms cost* OR economic* in more than one field.

The remaining three would be retrieved by a keyword search of: (*quality NEAR life*). (The search operator NEAR in this example means that the terms are within the same sentence.)

Therefore, a keyword search of *cost* OR economic** *OR* (*quality NEAR life*) would have retrieved all 91 economic evaluation studies used.

Delay in inclusion of studies in NHS EED

This was determined by analysing the 28 economic evaluation articles that were not in NHS EED at the time the searching was done, but were later indexed in NHS EED in 2002, Issue 4, of the Cochrane Library. Such an analysis can only give a minimum estimate of the delay, because other studies may still appear in the NHS EED at a later date.

Breakdown of the delay period for the 28 articles

- 7 = 0 years
- 10 = 1 year
- 4 = 2 years
- 3 = 3 years
- 2 = 4 years
- 2 = 5 years.

The two that took at least 5 years were from the journals *Anticancer Drugs* and *JAMA*, and the two that took 4 years were from *Journal of Pediatrics and Child Health* and *JAMA*.

Sources of studies used for economic modelling

In total, there were 121 articles cited in the section on economic modelling in 14 reviews. The articles were classified into the five different publication types: published, abstracts, reference sources, unpublished and grey literature. The relative proportions of each type, and their sources, are summarised below.

- Published = 50.4% (61). Fifty-eight (95.1%) of these were in MEDLINE. The three not in MEDLINE comprised two from EMBASE and one that could not be verified (probably an incorrect citation).
- Abstracts = 5.0% (6). Three of these were in SCI.
- Reference sources = 17.4% (21). Eight used government statistical sources, seven used data from *Unit Costs of Health and Social Care* and four cited the *British National Formulary (BNF)*.
- Unpublished = 17.4% (21). These were mostly obtained via personal communication and contact with experts.
- Grey literature = 9.9% (12). These included five websites, three NICE guidances and two DEC reports.

Chapter 5

Discussion and conclusions

Summary of main findings

Searching

- Reviews of cancer drugs were the most common topics of TARs in this sample.
- A total of 67 different sources was searched for the clinical and cost-effectiveness sections over the 20 TARs.
- The mean number of sources searched per TAR was 21 and the median was 20. The range was 13 to 33.
- Six sources (CCTR, DARE, EMBASE, MEDLINE, NHS EED and sponsor/industry submissions to NICE) were used in all reviews. CDSR and NRR were each searched in 18 TARs, and the HTA database and SCI were searched in 17 and 16 TARs, respectively. All ten of these sources were used in 15 (75%) of the TARs.
- Forty per cent of reviews stated that they excluded non-English language studies.
- Thirty-five per cent of reviews did not report whether the searching was restricted by language.

Clinical effectiveness studies

- Fifty-five per cent of TARs used evidence only from RCTs and systematic reviews; hence, 45% of TARs used some evidence from non-RCT studies (which include a wide range of study designs).
- Seventy-two per cent of the 424 studies were RCTs and 28% were non-RCTs.
- The publication types were: published 80%, meeting abstracts 11.3% and unpublished 8.7%.
- Eighty per cent of reviews included at least one abstract or unpublished study (60% included at least one abstract; 50% included at least one unpublished study).
- The mean number of studies used per TAR was 21.2 (median = 19.5; range 2 to 41).
- The mean number of participants included per TAR was 10,198 (median = 2787; range 69 to 97,570).
- Overall, 66% of participants were from RCTs and 34% from non-RCT studies.
- The median age of the studies used was 2 years.
- The percentage of studies classified as published and found indexed in the following databases was: MEDLINE 94.4%, EMBASE 89.7% and CCTR 55.8%.
- The percentage of studies classified as either published or abstracts and found indexed in the following databases was: MEDLINE 82.7%, EMBASE 78.6% and CCTR 50.1%.
- The cumulative percentage of studies classified as either published or abstracts and found indexed after searching three databases (MEDLINE, EMBASE and CCTR) was 87.3%. Adding SCI, BIOSIS and ASCO Online increased this to 98.2%.
- The overlap between MEDLINE and EMBASE was 87.4%.
- Six per cent (19) of the published studies were not in MEDLINE; 12 of the 19 were RCTs.
- None of the 12 non-Medline RCTs reported on the methods of allocation concealment used; four reported that the trial was double-blind.

Cost-effectiveness studies

Studies used in the review of economic evaluation

- The 130 articles were classified as: published 73.1%, unpublished 23.8% (nearly all from industry submissions), abstracts 1.5% and grey literature 1.5%.
- The median number of studies used was four, and 17 TARs found studies to use in the economic evaluation.
- The percentage of studies classified as either published or abstracts and found indexed in the following databases was: MEDLINE 86.6%, EMBASE 86.6% and NHS EED 40.2%.
- The cumulative percentage of studies classified as either published or abstracts, and found indexed after searching three databases (MEDLINE, EMBASE and NHS EED), was 94.8%. Adding SCI, and ASCO Online increased this to 97.9%.
- The overlap between MEDLINE and EMBASE was 86.7%.

Studies used in the economic modelling

- The articles were classified as: published 50.4%, abstracts 5.0%, reference sources 17.4%, unpublished 17.4% and grey literature 9.8%.
- The median number of studies used for the economic modelling was nine; 14 TARs included an economic model.

Search filters

- A sensitive filter for finding non-RCT studies in MEDLINE was devised. It was based on the search terms from the bibliographic records in the included studies only, and retrieved 85% of the sample.
- A simple search strategy for finding economic evaluations was devised. It was based on the search terms in the bibliographic records of the included economic evaluations, and retrieved 100% of studies in the sample.

Strengths of this study

To the authors' knowledge, this is the first study that has looked specifically at searching for TARs. The study analyses recent TARS and reviews of a range of interventions that have been selected so as to include some non-drug reports. It also covers clinical and cost-effectiveness sources, and reviews that include both RCT and non-RCT evidence. Finally, all studies were checked to see whether they were actually indexed in each of several databases, as opposed to just checking that the journal is indexed by the database. As some journals are selectively indexed, this method can be subject to inaccuracy. Although the majority of journals are indexed 'cover to cover', there are some journals where only selected articles that are thought relevant to biomedicine are indexed.

Weaknesses of this study

The number of TARs studied (20) was small, and there was a possible bias because of the high number of cancer topics in the NICE programme. Most of the key measurement decisions (checking of RCTs, classification of publication types and selection of potential search terms) were made by only one reviewer. In addition, the descriptions of the study designs used were as reported in the TARs and were not checked (the only exception to this was for those studies described as RCTs and not included in CCTR). It was also assumed that all TARs completely and accurately reported the sources that they searched.

This study did not look at the size and direction of treatment outcomes of the studies; hence, it was not possible to check whether the harder to find studies differed systematically in the size and direction of their treatment outcomes, and the impact that their exclusion would have made on the final results of any review. Neither did it evaluate the search terms that were used in the TARs, or look at the costs of searching the various sources. The section on devising search filters required some subjective judgement by the first author as to which search terms were considered most useful to describe the study design. In addition, because some of the commercial in confidence data had been removed from public versions of the TARs, the authors did not have access to the details of all studies that had been used in the reports.

Finally, the breadth of the NICE programme may be changing. Recent waves have included topics such as diabetes education and parenting for conduct disorders. While most TARs have been very focused, making searches easier, some of the more recent topics (not included in this study) have required a different approach to searching.

Discussion of key findings

Cost-effectiveness of literature searching

This study shows that there are diminishing marginal returns from increasingly exhaustive searching beyond a small number of major databases. Given the tight timescale and restricted resources for the TARS, one conclusion is that time could be better spent than on searching more and more databases. For example, it is probably more productive to spend time on checking the reference lists of retrieved studies, or asking experts to check lists for any study that may have been missed.

Those who would argue for exhaustiveness might assert that the last few studies retrieved from the more obscure sources might be in some way different, and might contribute more to the review than their numbers suggest. It is possible, although no evidence is available to support this, that studies with results that run counter to the conventional wisdom might have more difficulty being published, and thus might appear in less well-known journals which might not be indexed in the major databases.

However, one of the key findings of this study is that it is very rare for exhaustive database searching to find any studies that were used in TARs, which were not already found in the major databases. So, any argument about the last few studies found only in non-major databases being in some way different falls, because in practice it is rare to find any studies used that are not in the major databases. As *Table 12* shows, only seven out

of 424 studies published either in full or as an abstract could not be found in the six databases. Four of these were not indexed anywhere, and were probably identified through checking reference lists of studies that were retrieved, and a fifth was known to have been located in that way. One of the two others was indexed in PsycINFO, which would usually be added for mental health topics, and the last (a CRD report) could have been found in the Cochrane Library, in the HTA database.

Topics reviewed

The sample of TARs studied here had a predominance of drug interventions, especially cancer drugs. This is partly because NICE dealt with a batch of cancer drugs around this time, but it also reflects the prominence of both drugs and cancer in the NICE programme. Of the first 52 technology guidances issued, 35 were on drugs. Of the 15 that concerned cancer, all but two of these were on drugs.[33] Therefore, the sample reflects the content of the programme.

Difference in number of sources searched per review

There was quite a large variation between TARs in the number of sources searched (from 13 to 33). It is possible that some of this variation could be due to the differences between the different teams doing TARs in their searching practices. It was surprising that two reviews did not report that they searched the CDSR and three did not search the HTA database. Perhaps they had searched them very early on (maybe at the scoping stage), but did not report it in the TAR, especially if the intervention was very new and it was far too soon for a review to have been conducted.

The median number of sources searched for the TARs was 20 and the IQR was 16 to 27. This is considerably higher than for a sample of Cochrane reviews, where the median was 6 and the IQR was 5 to 9.[27] Possible reasons could be that TARs include searches for the cost-effectiveness sections (which Cochrane reviews normally do not), or the searchers' desire not to miss anything owing to the attention that TARs receive.

Language restrictions of studies used in TARs

Despite the fact that four reviews specifically mentioned that they did not restrict their searching by language (and another seven did not report whether there were any restrictions, so possibly none were imposed), only one non-English study (French) was used. It could be that no foreign language studies were relevant to the topics here, or none was of adequate quality to be included. Alternatively, this may reflect the number of studies that are carried out in the USA (because of the size of the market and of the research community) and hence published in American journals. Another reason may be that English has become the international language of science; for example, many of the main Scandinavian journals are published in English, such as the *Danish Medical Bulletin* and the *Journal of Internal Medicine* (formerly *Acta Medica Scandinavica*).

Eight reviews stated that they excluded non-English studies (although four of these eight identified the non-English studies, but did not include them in the review). This could be a potential source of bias in reviews. However, the influence of trials published in languages other than English on meta-analyses has been examined. It was found that non-English language trials included fewer participants and were more likely to produce significant results, and the methodological quality tended to be lower than that of trials published in English.[34] Estimates of treatment effects were found to be on average 16% more beneficial in non-English language trials than in English language trials, but the majority of meta-analyses excluding reports published in other languages did not change estimates of treatment effects substantially.

Nevertheless, it may be worth initially running the search strategies without the English language restriction to see the volume of research in non-English studies. Even though there would not be the resources to translate the articles, at least one would be aware of the volume of literature that was being excluded. In addition, as in most cases English abstracts for foreign language articles are available in the databases, this would allow one to gain some idea of whether there was a tendency for these to differ from the English language studies in the outcome and direction of the results.

Differences between TARs in the number of included studies

There was a large variation among the different TARs with respect to the total number of studies used in the clinical effectiveness section, with a maximum of 41 for asthma inhaler devices, down to as few as two for etanercept for juvenile idiopathic arthritis. This difference was also evident in the economic evaluations, with one review including 26 studies and two reviews not

identifying any studies. This variation would presumably be determined by many factors, including the type, scope and numbers of interventions, how new they are, how many comparators there are and the study designs to be included. For example, the review of lung cancer drugs included four drugs, one of which had 13 trials. Conversely, the TAR on temozolomide for brain cancer found only one RCT. The workload variation among different TARs reflects the number of studies.

TARs versus Cochrane reviews in the number of included studies

TARs that do not find RCTs or CCTs have to go down the evidence hierarchy and make use of the best available evidence; hence, a TAR with no included studies will not be produced. This reflects the need to present all the available evidence to policy makers, even if only to support a decision not to approve the technology until further research has been done. By contrast, some Cochrane reviews can have no included studies; that is, if no trials are found, they do not look for other evidence. (For example, it was found from a sample of Cochrane reviews that were added new to Issue 1 in 2001, and included RCTs only, that 12% did not find any trials that met the inclusion criteria.)[27] However, such reviews are useful in identifying areas where further research is needed.

Restriction by publication type

Four reviews stated that they excluded items published as abstracts only, and one review stated that it excluded abstracts for the economic evaluation searching. The danger of excluding abstracts is that one can potentially miss about 45% of the evidence, and risk overestimating treatment effects.[19] However, counteracting this is the fact that there are usually insufficient data to assess the quality of abstracts. This problem reflects the chronic tension between maintaining quality and striving to be as up to date as possible. Some TARs listed recent studies published as abstracts only in the appendices, but did not extract data from them, to give the reader an idea of research in progress.

The use of abstracts and unpublished data in TARs

It was found that 80.0% of the TARs used some data from either abstracts or unpublished data (60% included at least one abstract and 50% included at least one unpublished study). One reason for the frequent use of such data in TARs could be that many of them are identified from the industry submissions. It was noticeable that the three TARs that used the most abstracts were those on cancer drugs. The greater proportion of studies in abstract form may reflect fast-tracking by licensing authorities of cancer drugs, owing to the lack of currently effective drugs and the emotion that surrounds these diseases. Fast-tracking implies that the process of licensing is accelerated, which gives less time for studies to be published in full before NICE examines the new drug.

Searching for clinical effectiveness studies

In this study, the proportions of clinical effectiveness studies that were published either in full or as abstracts and found in the databases checked were: MEDLINE 82.7%, EMBASE 78.6% and CCTR 50.1%. As it was initially anticipated that most of the TARs would used RCTs, CCTR was selected as one of the three databases to be checked for studies included in the clinical effectiveness section. However, the much lower proportion of studies found in CCTR than expected is a reflection of the fact that about 28% of studies were non-RCTs and therefore would not be included in CCTR.

A search of these three databases gave a cumulative percentage of 87.3% of studies found. Adding SCI to the search yielded another 18 studies (4.7%), two of which were fully published articles and the remaining 16 were abstracts. Thus, it would seem that a search of SCI (after searching MEDLINE, EMBASE and CCTR) will retrieve some published studies not in the other three databases; although the yield is very small, it is occasionally worthwhile. (SCI indexes supplements to journals, some of which have full articles, that are not always indexed in MEDLINE and EMBASE.) However, the vast majority of extra studies retrieved in SCI will be meeting abstracts that appear in the journal supplements. A search of BIOSIS retrieved an extra four meetings abstracts. Finally, searching the ASCO database of abstracts retrieved 20 additional abstracts; but this step would only be appropriate for cancer reviews.

There were seven articles that were not retrieved (four published and three abstracts) after searching these six databases. A search of the HTA database on the Cochrane Library would have found one of them, one was in PsychINFO and another one was cited by another included study. Despite extensive searching, the remaining four were not found in any database. Therefore, it is likely that they were identified from the reference lists of other articles (perhaps the company submission) or by contact with experts in the

subject area. However, it is emphasised that this result only reflects the TAR topics studied here and the relative contributions of the different databases could be quite different for other topics.

As this is a retrospective study, it is not possible to know where the searchers identified their articles from; for example, just because they included a study that was indexed in MEDLINE does not necessarily mean that is where the searchers actually identified it. It is sufficient for the present purposes to know that the study could be found in MEDLINE. This study did not test the search strategies to determine from which database they would have retrieved the article.

The same search run on different databases will often retrieve different records, owing to the variation in search interfaces and indexing practices; therefore, sometimes something missed in one search will be picked up by searching another database. Consequently, there may be benefits to searching more than one database. This is probably not as important when the search is on a clearly defined drug name, but will be more so for a search topic that is more complex (e.g. a surgical or diagnostic technique, or a psychological intervention) and where the terminology is not as well defined.

The 'added value' of CCTR

There was only one record (an abstract in *Thorax*) that was in CCTR and would not have been found after searching MEDLINE, EMBASE, SCI, BIOSIS and ASCO Online. Although CCTR did not contribute very much in terms of the percentage of extra studies found, it is valuable as a check on the completeness of the searches of MEDLINE and EMBASE. This is because one can do a much more sensitive search for trials in CCTR, owing to the fact that it is a smaller database and is 'prefiltered' to include only CCTs; therefore, one can search on just the intervention alone.

Therefore, although nearly all of the trials found in CCTR were also indexed in at least one of the other main databases, they may have been missed when searching them, and actually only identified from the more sensitive search possible in CCTR. CCTR is very useful when doing scoping searches for reviews, and/or when one wants to estimate quickly the amount of good-quality evidence that exists on a particular topic without having to run complicated searches. However, there is a trade-off, as sensitive searches will require less searching time but more time for filtering of titles and abstracts.

The reliability of CCTR for finding RCTs

This study indicated that CCTR missed very few trials from MEDLINE, EMBASE and SCI, but a few do slip through. This, plus the fact that there is a delay for trials to be included in CCTR, indicates that it is worthwhile supplementing the CCTR search with a search of these three databases (at least for the past 2 years) using a fairly specific search filter for RCTs, particularly for the recent trials.

Analysis of published studies not indexed in MEDLINE

It was found that, with the exception of one study from *Current Therapeutic Research – Clinical and Experimental*, published trials not indexed in MEDLINE did not contribute very much to the total sample size. It was also found that the trials were of low quality, using the criteria of allocation concealment and double-blinding as measures of trial quality. However, the important questions not answered in this study were: (1) Are these trials of poorer quality relative to the other trials in the review? (2) How do they differ in terms of size and direction of treatment effects from the MEDLINE studies? (3) What would be the difference to the overall result if they were not included?

The use of non-RCT evidence in the TARs

This study found that 45% of the TARs used some evidence from non-RCTs in the clinical effectiveness section. There could be several reasons for this. One is that there may be few or no RCTs. Another is that although it is generally accepted that RCTs are the most reliable form of evidence in terms of internal validity, reviewers may sometimes look for observational, preferably population-based studies, as a check on the generalisability of the RCTs, and on whether the results achieved in trials are seen in routine care.

If non-RCT studies are to be included in a TAR this will usually increase the amount of time spent on the searching, because there are no reliable and tested filters and trial registers for such studies (as there are for RCTs). Normally, one needs to do a much more sensitive (and time-consuming) search on the intervention alone, and then scan all the abstracts for likely studies. In addition, because the study designs used in non-RCT studies are often not as well described as RCTs are in an abstract, reviewers often need to obtain the full paper to be able to determine the study design. This also adds to the time and expense of the review.

Search terms for identifying non-RCTs

A very sensitive filter constructed for MEDLINE, using the search terms from the bibliographic records in the included studies only, managed to retrieve 85% of the known sample. On this basis, it would seem unreliable to use a methodological filter for retrieving non-RCT studies, as certain relevant articles will be missed. Therefore, it is recommended that when searching for non-RCT studies a search is done just for the intervention alone, and records are then scanned manually for those that look relevant.

Searching for economic evaluations in different databases

The percentages of studies classified as published either in full or as abstracts, and found indexed in the following databases, were: MEDLINE 86.6%, EMBASE 86.6% and NHS EED 40.2%. A search of the three databases gave a cumulative percentage of 94.8%. Adding SCI gave one extra study and two additional abstracts were found from the ASCO database. The total percentage retrieved from these five databases was 97.9%. There were two articles that were not found and one was cited by another included article. It is likely that the remaining article was cited by one of the clinical effectiveness articles, as they both had an author in common.

The publication types of the 130 economic evaluation articles were: published 95 (73.1%), unpublished 31 (23.8%), abstracts 2 (1.5%) and grey literature 2 (1.5%). Thus, a much higher percentage of unpublished studies (nearly all from company submissions) was used for the economic evaluations, compared with the 8.7% of the clinical effectiveness studies that were unpublished. This is probably because only clinical effectiveness evidence has been required for licensing, whereas NICE bases its decisions on both clinical effectiveness and cost-effectiveness. There is often no published study of cost-effectiveness for an intervention, or there may be studies conducted in other countries and so not immediately applicable to the UK. The manufacturers may therefore have to commission new work, often using unpublished data from trials, but tailored to NHS use of the product. However, perhaps the main reason is that the manufacturers will seek to convince NICE of the cost-effectiveness of their product, and will tend to commission a favourable report that fits their purpose.

It is likely that there is more variation in the approach to searching for cost-effectiveness compared with clinical effectiveness, as although there are many guides on how to do the searching for the effectiveness part of the review, this is not the case in areas such as cost-effectiveness, decision modelling and adverse effects. A simple search strategy (based on terms occurring in all of the included economic evaluations) was presented in this study for retrieving economic analyses, but this was based only on a small sample, and more sophisticated and sensitive strategies are recommended to ensure a more thorough search.[35–37]

A recent study by Sassi and colleagues[37] on searching literature databases for healthcare economic evaluations concluded that researchers may confidently rely on MEDLINE as the key source for the identification of published economic evaluations, and that searches of other electronic literature databases and manual searches of journals and grey literature appear to provide a limited additional yield.

The 'added value' of NHS EED

This is similar to the added value for CCTR (see above), in that although it did not appear to add much in terms of the proportion of unique studies identified, it is a useful check on the completeness of the searching for economic evaluations. As NHS EED is a relatively small database, one can do a very sensitive search for economic evaluations by just searching on the terms for the intervention alone, without having to use a search filter. It is noted that NHS EED is prospective from 1995 onwards, so the pre-1995 literature has to be identified from elsewhere.

Only two studies found in NHS EED were not also found in MEDLINE, EMBASE or SCI. One was published in *Applied Economics* (which is indexed in SSCI) and one in the grey literature – a *HERU Discussion Paper*. Therefore, the contribution of NHS EED was low in terms of the quantity of new studies found identified for the reviews in this sample, but its added value is due to the structured abstracts available. These are intended to provide searchers with rapid and comprehensive details of the original papers, to help them to decide whether papers are relevant and of sufficient quality to be of use in their decision-making processes.

Unfortunately, the authors did not have access to the Health Economic Evaluations Database (HEED); it would have been interesting to see how it compared with NHS EED.

Searching for sources to use in the economic modelling

The proportions of the different publication types cited in the economic modelling sections were:

published 50.4% (61), abstracts 5.0% (6), reference sources 17.4% (21) and unpublished 17.4% (21). The unpublished sources cited were mainly personal communications, and the most commonly used reference sources were BNF, Unit Costs of Health and Social Care, and government statistical sources.

There was a tendency for those reports that cite a lot of economic evaluations to cite fewer sources for the modelling, and vice versa. Therefore, it may be that if there is a substantial body of existing economic literature, there is less need for modelling: or, alternatively, that a need for modelling sometimes reflects a lack of data.

The data required for economic modelling vary according to topic, but in addition to data from the clinical effectiveness review, are likely to include at least studies on:

- the natural history of the disease, if there is no current treatment, or the results of current best treatment if there is one, as a baseline for outcomes and costs without the new intervention. The information needed would include effects on mortality, morbidity and quality of life
- the prevalence or incidence of the disease, as a guide to possible costs to the NHS
- long-term results of the intervention that may reveal rare but serious side-effects not detectable in the trials, because of small numbers, patient mix or a short duration of follow-up. Some costly side-effects may become apparent only after several years
- results in routine care, which may not be as good as in the trials
- cost data, which may have to come from grey sources. Ideally, costs should come from more than one hospital, since costs, and ways of estimating them, may vary among hospitals
- acceptability to patients, and hence compliance and costs in routine care.

The overlap between MEDLINE and EMBASE

The overlap between MEDLINE and EMBASE has been variously estimated at between 10 and 79%, depending on the topic being searched and the methods used to measure it.[26,27,38] The overlap between MEDLINE and EMBASE in this study was found to be 87%. Hence, it would currently still seem worth searching both databases.

One reason that the overlap could be higher in this study is because the articles used here are more recent, and both MEDLINE and EMBASE are steadily increasing the number of journal titles that they index. For example, it was found in this study that three journals that were unique to EMBASE in the late 1990s have since been picked up by MEDLINE, and five journals that were unique to MEDLINE are now in EMBASE. If the trend for these databases to converge in their journal coverage continues, then theoretically there will eventually be only the need to search one database. However, the two databases have different indexing practices so, as mentioned previously, an article indexed in both databases may not always be retrieved from both when a similar search strategy is run.

Is EMBASE superior to MEDLINE in the coverage of drug journals?

EMBASE is commonly thought as being superior to MEDLINE when searching for drug studies.[39] However, this study found that only two of a total of seven 'EMBASE-only' journals (i.e. those not also in MEDLINE) were exclusively pharmaceutical journals. In a study of Cochrane reviews it was found that of the 32 EMBASE-only journals, only five journals were pharmaceutical/drug journals.[40] These observations run contrary to the common perception of EMBASE being particularly useful in drug reviews. However, perhaps the fact that this study did not confirm this perception merely reflects the subjects looked at in the reviews studied, and EMBASE may have better coverage of other pharmaceutical journals not used in these reviews.

Time-lag bias

Time-lag bias is defined as the rapid or delayed publication of research findings, depending on the nature and direction of the results.[12] It was found in this study that some TARs use recently published data for the clinical effectiveness evidence (the median age of studies was 2 years and the IQR was 1 to 6 years). This use of recent studies presents the danger of time-lag bias in reviews, especially in those reviews that include only a few recent studies.

Evidence for time-lag bias comes from a recent systematic review that investigated the extent to which the time to publication of a clinical trial is influenced by the significance of its result.[41] The study showed that trials with null or negative findings took, on average, just over 1 year longer to be published than those with positive results. The authors noted the importance of the implications that this has for the timing of the initiation and updating of a review, especially if only very few studies are currently available.

The role of commercial in confidence data in TARs

Currently, two versions of TARs are produced: one containing the commercial in confidence data (which goes to the NICE appraisal committee) and a second version in the public domain, with the confidential data removed. This is problematic in terms of the reproducibility of the review, as it is possible that the conclusions will not always follow logically from the data presented in the public version.

The impact of access to manufacturers' submission for TARs

The manufacturers' submissions can be used in several ways. First, the submission will provide evidence from trials, and this can be used as a check on completeness of ascertainment of all RCTs, particularly as so many trials are funded by the manufacturers, but also because they will usually maintain databases of studies on their products. The TAR search may, however, have found trials not quoted by the manufacturer, perhaps because the results are not so favourable. Second, the manufacturer may provide data from unpublished trials. These may be 'academic in confidence' until the researchers have secured publication in a peer-reviewed journal. (Premature release of results might imperil publication.) Third, the manufacturer may have unpublished data from trials that have been published; journal space is often limited and it may not be possible to fit all of the data in.

However, the TAR team will bear in mind that the manufacturer has a vested interest, and may present selected evidence or a biased economic evaluation, and so the manufacturer's submission will be treated with some caution and may not carry much weight in the final TAR, compared with evidence from independent sources.

Costs of databases

The relative costs of databases can vary greatly and this can be an important factor when deciding on which ones to search. These costs were not investigated in this study, as the complex and varied pricing structures make it extremely difficult to compare costs between different database vendors, institutions and countries.

MEDLINE (via PubMed) and the CRD databases (DARE, NHS EED and HTA) are available free of charge. However, the Cochrane Library is currently only freely available in Australia, England, Finland, Ireland, Norway and Wales; users from other countries have to pay a subscription charge.

Most users in UK higher educational institutions (where TAR teams are based) have unlimited access to the Cochrane Library, EMBASE, SCI, Social Sciences Citation Index (SSCI) and PsycINFO (and other databases) through centrally funded services, so the costs involved in additional searching are purely labour costs. However, in North America and some parts of Europe, these databases are accessed via commercial vendors where the pricing structure includes costs for connect time and the downloading of records, so searching costs can be quite substantial. For this reason, many health researchers and professionals from these countries will rely on MEDLINE alone for searching.[26] Therefore, it is especially important that users are aware of the added value of searching any additional databases.

The tension between efficiency and making reports immune from criticism

The TARs produced for NICE are scrutinised extremely closely by the sponsors of technologies and by the consumer groups involved. This is because they are released to the consultees to the appraisal (including manufacturers, patient groups and clinical experts) before the topic is considered at the first meeting of the Appraisal Committee. If the TAR suggests that a new product is not cost-effective, the industrial sponsor may carry out or commission an intensive critique of the TAR, looking for flaws. This will include checking for any studies that have been missed or excluded, for example on quality grounds such as design or choice of comparator. There is, therefore, pressure on the TAR team to find every study, including those that are poor and would be excluded, since otherwise the absence of the study could be used as a criticism. A related issue is that poorer quality studies may show greater effect size, and may be particularly likely to be quoted. Assessment reports prepared for the NICE appraisal process may be subject to further detailed scrutiny should the appraisal policy decision be subject to appeal.

Hence, it is possible that unnecessary searches may be carried out in the same way as unnecessary X-rays may be taken in casualty departments – not for the good that they do, but as a form of 'defensive medicine'. However, this study provides further evidence that this is wasteful of scarce resources.

Comparison with the study by Egger and colleagues

The study by Egger and colleagues[9] was described in Chapter 2. Their study differed from this one in several key aspects.

- They aimed to examine the characteristics of clinical trials that were difficult to locate, and compare within meta-analyses the treatment effects reported in such trials with trials that are more accessible, and trials of lower with trials of higher quality; the aim of the present study was to look at bibliographic sources, i.e. where studies are found, not their impact on the conclusions of the review.
- They only included RCTs from meta-analyses that combined the binary outcomes of at least five controlled trials.
- They did not include any TARs, whereas this study included only TARs, which use a wider range of study designs than just RCTs.
- Their definition of 'difficult to find' trials was trials that were either unpublished, published in languages other than English or not indexed in MEDLINE. However, this definition did not include the CCTR. As this database (based not only on studies retrieved by searching electronic databases, but also by the handsearching of journals carried out by members of Collaborative Groups) continues to increase, it will greatly reduce the number of trials that can be described as difficult to locate. It should be noted that CCTR already contains a considerable proportion of trials from the grey literature, which formerly would have been time consuming and difficult to find.

Despite the different aims and methods, the overall messages are similar – that exhaustive searching of databases to find difficult to locate trials is usually not worthwhile. The authors cannot comment on the impact on conclusions of studies not found in the major databases, because too few were found for any quality comparison.

Research needs

There is a need for prospective studies, over a wide range of topics, to investigate further the effectiveness of extended searches in identifying extra studies. One would need to measure the relative quality, and size and direction of outcome of the more difficult to find studies relative to those easily found ones. There is also a need for data on the costs involved in searching and acquiring these extra studies relative to the easily found ones. In addition, one would need to perform sensitivity analyses to measure the impact of these studies on the final results of a systematic review.

It would be useful to test the generalisability of the findings from this study in an international context by doing a comparative analysis on technology assessment reports produced by other agencies in the INAHTA.

However, one major problem in any statistical analysis involving such comparative studies is that the extra studies found by searching beyond the four major databases may be few in number, so that the statistical power of the comparisons could be too small to come to any valid conclusion.

It would appear that there is a need for some research into assessing the quality of search strategies used in systematic reviews. The quality of the search strategy may to some extent determine the number of sources needed to conduct a thorough search; for example, a very good strategy run on MEDLINE may reduce the need to search a lot of other databases. At present, there is no agreed method for determining the quality of a search strategy, so some way of doing this would first have to be devised.

One impression is that drug studies require fewer databases to be searched than some non-drug studies; perhaps drug journals are more likely to be indexed in major databases. Alternatively, it may be that because the drug names are precisely defined they are easier to retrieve in any search with a few simple keywords. It would be useful to test this impression prospectively on some drug reviews and non-drug reviews, noting which database identified each included study.

There is also a need for the development and testing of search filters for retrieving each of the many different types of non-RCT study. One would need to test the various strategies devised against a 'gold standard' of studies, and measure their specificity and sensitivity across a range of different subject areas.

Finally, it would also be interesting to conduct a follow-up study to see what proportion of the unpublished drug company studies are eventually published, and whether the conclusions of the published versions differ from the commercial in confidence versions. If the published versions are less favourable to the product, or indeed more favourable, the guidance may have suffered from bias. Any differences could also create problems for others seeking to replicate or update the systematic review, since a review of only evidence in the public domain might reach different conclusions. This may be a particular problem (although there is no evidence of this) if NICE is looking at products closer to launch, when a

smaller proportion of the evidence may have been published in full in peer-reviewed journals. The confidential material available to the Appraisal Committees includes not only commercial in confidence data but also academic in confidence material that is being submitted for publication by the researchers; release of data through the NICE processes might imperil publication in leading medical journals.

Recommendations and implications of this study for TAR teams

On the basis of data presented in this study it is recommended that the Cochrane Library, MEDLINE, EMBASE and SCI be searched in all reviews. It is worth also searching the CRD databases (DARE, NHS EED and the HTA database) on the web, as although they are included in the Cochrane Library, the web versions are more up to date. The HTA database is particularly useful for those undertaking technology assessments, as it contains records of ongoing projects and completed technology assessments conducted by other technology assessment organisations, most of which are not indexed in MEDLINE or EMBASE. A search of PubMed, limited to records added in the past 90 days and using free text words, is also recommended to retrieve very recent articles that have not yet been fully indexed in MEDLINE.

The ASCO database should also be searched when doing cancer reviews. If meeting abstracts are to be included in a review, then one could add a search of BIOSIS, but limit it to meeting abstracts only. It is also useful to check the websites of any relevant professional organisations, as many make available the abstracts from their recent conferences, and these appear some months before they are published in the journal supplements. For example, the website of the American Diabetes Association can contain the abstracts of their conferences some months before they appear in the supplement to *Diabetes*.

It is also recommended that NRR, Controlled Trials.com, and ClinicalTrials.gov (all freely available on the web) are searched to check for recently completed and ongoing searches, and a handsearch of recent issues of relevant journals that have not yet been indexed in the major databases. In addition, as it has been shown that the FDA can be an important source of unpublished trials,[20,26] it would seem useful to check the FDA website when doing a review. Awareness of ongoing research is useful for scheduling the next review of evidence.

When conducting searches for economic evaluations it could also be useful to include SSCI, as there are economics journals that are indexed in SSCI, but not in SCI. However, further study of the usefulness of SSCI is needed. It may also be worth searching PsycINFO and ERIC, for topics involving psychological therapies and educational interventions, respectively, although the usefulness of these sources in TARs has yet to be determined.

The time saved on extra database searching could be spent on other methods of obtaining potentially relevant studies, such as checking recent meeting abstracts, contacting experts for recently completed studies and unpublished data, and scrutinising the references provided in the industry submissions.

It was found in this study that there was a good deal of variation between TARs in the way in which search strategies were reported and in the amount of detail given (e.g. reports varied in how the search terms used were presented, whether the restrictions on searching by language were reported, and whether the versions of databases used were given). Perhaps the teams doing TARs could agree to standardise the reporting of the search strategies, to make the reporting more uniform and to ensure that the appropriate amount of detail is given.

Finally, it should be recognised that any search strategy needs to be adapted to the particular topic under review. Although drug trials tend to be concentrated in mainstream journals and are therefore widely indexed, more diverse searches may be required for non-drug or educational/psychological interventions. The findings from this study cannot necessarily be extrapolated to all types of technology assessment, as the reviews studied here reflect the agenda of NICE over a period when it was dominated by assessments of new drugs. The range of NICE guidance is likely to change over time.

The impression gained is that hard to find trials tend to be of poorer quality, so may be rapidly discarded after perusal. However, because the number of trials found only in the less commonly used databases is very low, it is not easy to prove the quality difference statistically.

A more efficient approach, using the four major databases for clinical effectiveness, will save resources by reducing:

- licence fees, online search time costs and costs of acquiring grey literature
- information specialists' time for searching the extra databases and filtering the product
- reviewers' time for checking (or double checking) abstracts and full papers.

In the course of the review, it was noted that search practices vary among the academic teams that produce the TARs. The findings of this study, and of other work such as that by Egger and colleagues, will help discussions on standardising search strategies.

Conclusions

Searching additional databases beyond the Cochrane Library, MEDLINE, EMBASE and SCI, plus BIOSIS limited to meeting abstracts only, is seldom effective in retrieving additional studies for inclusion in the clinical and cost-effectiveness sections of TARs (apart from reviews of cancer therapies, where a search of the ASCO database is recommended). A more selective approach to database searching would suffice in most cases and would save resources, thereby making the TAR process more efficient. However, searching of non-database sources (including submissions from manufacturers, recent meeting abstracts, contact with experts and checking reference lists) does appear to be a productive way of identifying further studies.

There may be a tension between the cost-effectiveness of searching and the desire to protect the reviews from criticism by being seen to have searched everything. However, this study adds to the body of evidence that exhaustive database searching usually adds little, and provides further protection against any criticism. It is likely that more exhaustive searches of an extended number of databases have been carried out more for reassurance and as a matter of procedure, than for any extra evidence of use to the TAR.

Acknowledgements

This report was produced by the Department of Public Health, University of Aberdeen and Southampton Health Technology Assessment Centre, University of Southampton. Some of this work was done while the authors were working at the University of Southampton.

We are grateful to the following for expert advice and comments on the draft report, but we absolve them from any faults in the final version: Suzy Paisley, Managing Director, ScHARR; Julie Glanville, Associate Director, Centre for Reviews and Dissemination; Ruth Frankish, Senior Information Specialist, Appraisals, NICE; and Leigh-Ann Topfer, Research Officer, Canadian Coordinating Office for Health Technology Assessment (CCOHTA).

Contributions of authors

Pamela Royle conceived the study, co-wrote the protocol, extracted and managed the data, did the literature searching and co-wrote the report. Norman Waugh co-wrote the protocol and the report.

References

1. Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987;**106**:485–8.
2. Clarke M, Oxman AD, editors. Cochrane Reviewer's Handbook 4.1.6 (updated January 2003). http://www.cochrane.dk/cochrane/handbook/handbook.htm. Accessed 23 April 2003.
3. Glasziou P, Irwig L, Bain C, Colditz G. Systematic reviews in health care: a practical guide. Cambridge: Cambridge University Press; 2001.
4. Cochrane Collaboration. http://www.cochrane.de
5. Antes G, Oxman AD. The Cochrane Collaboration in the 20th century. In Egger M, Davey SG, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. London: BMJ Books; 2001. pp. 447–73.
6. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, *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**:278–80.
7. National Institute for Clinical Excellence. http://www.nice.org.uk
8. Best L, Stevens A, Colin Jones D. Rapid and responsive health technology assessment: the development and evaluation process in the South and West region of England. *Journal Clinical Efficiency* 1997;**2**:51–6.
9. Egger M, Juni P, Bartlett C, Holenstein F, Sterne J. How important are comprehensive literature searches and the assessment of trial quality in systematic reviews? Empirical study. *Health Technol Assess* 2003;**7**:1–76.
10. Egger M, Davey SG, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. London: BMJ Books; 2001.
11. Song F, Eastwood AJ, Gilbody S, Duley L, Sutton AJ. Publication and related biases. *Health Technol Assess* 2000;**4**(10):1–115.
12. Egger M, Dickersin K, Davey SG. Problems and limitations in conducting systematic reviews. In Egger M, Davey SG, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. London: BMJ Books; 2001. pp. 43–68.
13. Campanario JM. Consolation for the scientist – sometimes it is hard to publish papers that are later highly-cited. *Social Studies of Science* 1993;**23**:342–62.
14. Spier RE. Peer review and innovation. *Sci Eng Ethics* 2002;**8**:99–108.
15. Pittler MH, Abbot NC, Harkness EF, Ernst E. Location bias in controlled clinical trials of complementary/alternative therapies. *J Clin Epidemiol* 2000;**53**:485–9.
16. Lee KP, Schotland M, Bacchetti P, Bero LA. Association of journal quality indicators with methodological quality of clinical research articles. *JAMA* 2002;**287**:2805–8.
17. Hill DR, Stickell HN. Brandon/Hill selected list of print books and journals for the small medical library. *Bulletin of the Medical Libraries Association* 2001;**89**:131–53.
18. Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions (Cochrane Methodology Review). In The Cochrane Library (Issue 1). Oxford: Update Software; 2003.
19. Scherer RW, Langenberg P. Full publication of results initially presented in abstracts (Cochrane Methodology Review). Cochrane Library. Oxford: Update Software; 2001.
20. MacLean CH, Morton SC, Ofman JJ, Roth EA, Shekelle PG. How useful are unpublished data from the Food and Drug Administration in meta-analysis? *J Clin Epidemiol* 2003;**56**:44–51.
21. Cook DJ, Guyatt GH, Ryan G, Clifton J, Buckingham L, Willan A, *et al*. Should unpublished data be included in meta-analyses? Current convictions and controversies. *JAMA* 1993; **269**:2749-53.
22. Jefferson TO, Alderson P, Davidoff F, Wager E. Editorial peer-review for improving the quality of reports of biomedical studies (Cochrane Methodology Review). In The Cochrane Library (Issue 1). Oxford: Update Software; 2003.
23. Savoie I, Helmer D, Green CJ, Kazanjian A. Beyond Medline: reducing bias through extended systematic review search. *Int J Technol Assess Health Care* 2003;**19**:168–78.
24. Sampson M, Campbell K, Ajiferuke I, Moher D. Randomized controlled trials in pediatric complementary and alternative medicine: where can they be found? *BMC Pediatr* 2003;**3**:1.
25. Avenell A, Handoll HHG, Grant AM. Lessons for search strategies from a systematic review, in The Cochrane Library, of nutritional supplementation trials in patients after hip fracture. *Am J Clin Nutr* 2001;**73**:505–10.

26. 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. *Control Clin Trials* 2000;**21**:476–87.

27. Royle P, Milne R. Literature searching for randomised controlled trials used in Cochrane reviews: rapid versus exhaustive searches. *Int J Technol Assess Health Care* 2003; in press.

28. Cook AM, Finlay IG, Edwards AG, Hood K, Higginson IJ, Goodwin DM, *et al*. Efficiency of searching the grey literature in palliative care. *J Pain Symptom Manage* 2001;**22**:797–801.

29. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, *et al*. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.

30. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *BMJ* 1996;**313**:275–83.

31. Lefebvre C, Clarke MJ. Identifying randomised trials. In Egger M, Davey SG, Altman DG, editors. Systematic reviews in health care: meta-analysis in context. London: BMJ Books; 2001. pp. 69–86.

32. Egger M, Ebrahim S, Smith GD. Where now for meta-analysis? *Int J Epidemiol* 2002;**31**:1–5.

33. NICE. Compilation – summary of guidance issued to the NHS in England and Wales. London: National Institute for Clinical Excellence; October 2002 (Issue 5).

34. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. *Int J Epidemiol* 2002;**31**:115–23.

35. NHS Centre for Reviews and Dissemination. Centre for Reviews and Dissemination (CRD). Improving access to cost-effectiveness information for health care decision making: the NHS Economic Evaluation Database. CRD Report 6. 2nd ed. York: University of York; 2001.

36. Eady A. Secondary publications: economic analyses. In McKibbon A, editor. PDQ evidence-based principles and practice. Hamilton: Decker; 1999. pp. 173–86.

37. Sassi F, Archard L, McDaid D. Searching literature databases for health care economic evaluations: how systematic can we afford to be? *Med Care* 2002; **40**:387–94.

38. Clarke M, Oxman AD, editors. Locating and selecting studies. Cochrane Reviewers Handbook 4.1.6 (updated January 2003) Section 5. In The Cochrane Library (Issue 2). Oxford: Update Software; 2003.

39. Topfer LA, Parada A, Menon D, Noorani H, Perras C, Serra-Prat M. 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.

40. Royle P. Literature searching for randomised controlled trials used in Cochrane reviews: rapid versus exhaustive searches [dissertation]. Southampton: University of Southampton; 2001.

41. Hopewell S, Clarke M, Stewart L, Tierney J. Time to publication for results of clinical trials (Cochrane Methodology Review). In The Cochrane Library (Issue 1). Oxford: Update Software; 2003.

Appendix

Sources searched in the TARS

Name of source	No. of reviews in which it was searched
CCTR	20
DARE	20
EMBASE	20
MEDLINE	20
NHS EED	20
Sponsor/industry submissions to NICE	20
CDSR	18
NRR	18
HTA database	17
SCI	16
Web searching	15
Reference lists	13
BIOSIS/Biological Abstracts	12
PubMED/PreMedline	12
Controlled Trials.com	10
CINAHL	9
Experts	9
ISTP	9
HEED (OHE EED)	8
HMIC	8
NIH Clinical Trials Register	7
SSCI	7
CenterWatch Clinical Trials	6
EconLit	6
Cancerlit	5
ClinicalTrials.gov	5
CRB	5
HealthSTAR	5
Manufacturers	5
MRC Clinical Trials	5
PsycINFO/PsycLit	5
Websites – health economics related	5
AMED	4
ASCO database 1997–2000	4
Bandolier	4
ReFer-DH Research Findings Register	4
SIGN guidelines	4
Trip database	4
Cited reference searching with SCI	3
In-house databases	3
NCI web page	3
Abstracts of meetings	2
Best evidence	2
BNI	2
DH data	2
EBM reviews	2
HELMIS	2
Kings Fund Database	2
NLM Gateway	2
UKCCR	2
Audit databases	1
Authors contacted	1
BNF	1
British Library/Inside	1
CancerTrials	1
Clinical evidence	1
Conference Papers Index	1
CRW databases	1
Econbase	1
EWS	1
FDA website	1
Harvard Database of Cost–Utility Analysis	1
Martindale Pharmacopoeia	1
Physicians Data Query	1
Scrip	1
Toxline	1
Unpublished studies sought	1

ISTP: Index to Scientific and Technical Proceedings; OHE EED: Office of Health Economics Economic Evaluation Database; NIH: National Institutes of Health; CRB: Current Research in Britain; MRC: Medical Research Council; AMED: Allied and Complementary Medicine; DH: Department of Health; SIGN: Scottish Intercollegiate Guidelines Network; NCI: National Cancer Institute; EBM: Evidence Based Medicine; NLM: National Library of Medicine; UKCCR: UK Co-ordinated Committee or Cancer Research; CRW: Current Research Worldwide; EWS: Early Warning System.

continued

Health Technology Assessment Programme

Prioritisation Strategy Group

Members

Chair,
Professor Kent Woods,
Director, NHS HTA Programme & Professor of Therapeutics, University of Leicester

Professor Bruce Campbell, Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital

Professor Shah Ebrahim, Professor in Epidemiology of Ageing, University of Bristol

Dr John Reynolds, Clinical Director, Acute General Medicine SDU, Radcliffe Hospital, Oxford

Dr Ron Zimmern, Director, Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

HTA Commissioning Board

Members

Programme Director,
Professor Kent Woods, Director, NHS HTA Programme, Department of Medicine and Therapeutics, Leicester Royal Infirmary, Robert Kilpatrick Clinical Sciences Building, Leicester

Chair,
Professor Shah Ebrahim,
Professor in Epidemiology of Ageing, Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol

Deputy Chair,
Professor Jenny Hewison,
Professor of Health Care Psychology, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds School of Medicine, Leeds

Professor Douglas Altman, Professor of Statistics in Medicine, Centre for Statistics in Medicine, Oxford University, Institute of Health Sciences, Cancer Research UK Medical Statistics Group, Headington, Oxford

Professor John Bond, Professor of Health Services Research, Centre for Health Services Research, University of Newcastle, School of Health Sciences, Newcastle upon Tyne

Professor John Brazier, Director of Health Economics, Sheffield Health Economics Group, School of Health & Related Research, University of Sheffield, ScHARR Regent Court, Sheffield

Dr Andrew Briggs, Public Health Career Scientist, Health Economics Research Centre, University of Oxford, Institute of Health Sciences, Oxford

Dr Christine Clark, Medical Writer & Consultant Pharmacist, Cloudside, Rossendale, Lancs
and
Principal Research Fellow, Clinical Therapeutics in the School of Pharmacy, Bradford University, Bradford

Professor Nicky Cullum, Director of Centre for Evidence Based Nursing, Department of Health Sciences, University of York, Research Section, Seebohm Rowntree Building, Heslington, York

Dr Andrew Farmer, Senior Lecturer in General Practice, Department of Primary Health Care, University of Oxford, Institute of Health Sciences, Headington, Oxford

Professor Fiona J Gilbert, Professor of Radiology, Department of Radiology, University of Aberdeen, Lilian Sutton Building, Foresterhill, Aberdeen

Professor Adrian Grant, Director, Health Services Research Unit, University of Aberdeen, Drew Kay Wing, Polwarth Building, Foresterhill, Aberdeen

Professor Alastair Gray, Director, Health Economics Research Centre, University of Oxford, Institute of Health Sciences, Headington, Oxford

Professor Mark Haggard, Director, MRC ESS Team, CBU Elsworth House, Addenbrooke's Hospital, Cambridge

Professor F D Richard Hobbs, Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham, Primary Care and Clinical Sciences Building, Edgbaston, Birmingham

Professor Peter Jones, Head of Department, University Department of Psychiatry, University of Cambridge, Addenbrooke's Hospital, Cambridge

Professor Sallie Lamb, Research Professor in Physiotherapy/Co-Director, Interdisciplinary Research Centre in Health, Coventry University, Coventry

Dr Donna Lamping, Senior Lecturer, Health Services Research Unit, Public Health and Policy, London School of Hygiene and Tropical Medicine, London

Professor David Neal, Professor of Surgical Oncology, Oncology Centre, Addenbrooke's Hospital, Cambridge

Professor Tim Peters, Professor of Primary Care Health Services Research, Division of Primary Health Care, University of Bristol, Cotham House, Cotham Hill, Bristol

Professor Ian Roberts, Professor of Epidemiology & Public Health, Intervention Research Unit, London School of Hygiene and Tropical Medicine, London

Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh, Western General Hospital NHS Trust, Bramwell Dott Building, Edinburgh

Professor Martin Severs, Professor in Elderly Health Care, Portsmouth Institute of Medicine, Health & Social Care, St George's Building, Portsmouth

Dr Jonathan Shapiro, Senior Fellow, Health Services Management Centre, Park House, Birmingham

Current and past membership details of all HTA 'committees' are available from the HTA website (www.ncchta.org)

Diagnostic Technologies & Screening Panel

Members

Chair,
Dr Ron Zimmern, Director of the Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

Dr Paul Cockcroft, Consultant Medical Microbiologist/ Laboratory Director, Public Health Laboratory, St Mary's Hospital, Portsmouth

Professor Adrian K Dixon, Professor of Radiology, Addenbrooke's Hospital, Cambridge

Dr David Elliman, Consultant in Community Child Health, London

Dr Andrew Farmer, Senior Lecturer in General Practice, Institute of Health Sciences, University of Oxford

Dr Karen N Foster, Clinical Lecturer, Dept of General Practice & Primary Care, University of Aberdeen

Professor Jane Franklyn, Professor of Medicine, University of Birmingham

Professor Antony J Franks, Deputy Medical Director, The Leeds Teaching Hospitals NHS Trust

Mr Tam Fry, Honorary Chairman, Child Growth Foundation, London

Dr Susanne M Ludgate, Medical Director, Medical Devices Agency, London

Dr William Rosenberg, Senior Lecturer and Consultant in Medicine, University of Southampton

Dr Susan Schonfield, CPHM Specialised Services Commissioning, Croydon Primary Care Trust

Dr Margaret Somerville, Director of Public Health, Teignbridge Primary Care Trust, Devon

Mr Tony Tester, Chief Officer, South Bedfordshire Community Health Council, Luton

Dr Andrew Walker, Senior Lecturer in Health Economics, University of Glasgow

Professor Martin J Whittle, Head of Division of Reproductive & Child Health, University of Birmingham

Dr Dennis Wright, Consultant Biochemist & Clinical Director, Pathology & The Kennedy Galton Centre, Northwick Park & St Mark's Hospitals, Harrow

Pharmaceuticals Panel

Members

Chair,
Dr John Reynolds, Clinical Director, Acute General Medicine SDU, Oxford Radcliffe Hospital

Professor Tony Avery, Professor of Primary Health Care, University of Nottingham

Professor Iain T Cameron, Professor of Obstetrics & Gynaecology, University of Southampton

Mr Peter Cardy, Chief Executive, Macmillan Cancer Relief, London

Dr Christopher Cates, GP and Cochrane Editor, Bushey Health Centre, Bushey, Herts.

Mr Charles Dobson, Special Projects Adviser, Department of Health

Dr Robin Ferner, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham

Dr Karen A Fitzgerald, Pharmaceutical Adviser, Bro Taf Health Authority, Cardiff

Professor Alastair Gray, Professor of Health Economics, Institute of Health Sciences, University of Oxford

Mrs Sharon Hart, Managing Editor, *Drug & Therapeutics Bulletin*, London

Dr Christine Hine, Consultant in Public Health Medicine, Bristol South & West Primary Care Trust

Professor Robert Peveler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton

Dr Frances Rotblat, CPMP Delegate, Medicines Control Agency, London

Mrs Katrina Simister, New Products Manager, National Prescribing Centre, Liverpool

Dr Ken Stein, Senior Lecturer in Public Health, University of Exeter

Professor Terence Stephenson, Professor of Child Health, University of Nottingham

Dr Richard Tiner, Medical Director, Association of the British Pharmaceutical Industry, London

Professor Dame Jenifer Wilson-Barnett, Head of Florence Nightingale School of Nursing & Midwifery, King's College, London

Therapeutic Procedures Panel

Members

Chair,
Professor Bruce Campbell,
Consultant Vascular and General Surgeon, Royal Devon & Exeter Hospital

Dr Mahmood Adil, Head of Clinical Support & Health Protection, Directorate of Health and Social Care (North), Department of Health, Manchester

Professor John Bond, Head of Centre for Health Services Research, University of Newcastle upon Tyne

Mr Michael Clancy, Consultant in A & E Medicine, Southampton General Hospital

Dr Carl E Counsell, Senior Lecturer in Neurology, University of Aberdeen

Dr Keith Dodd, Consultant Paediatrician, Derbyshire Children's Hospital, Derby

Professor Gene Feder, Professor of Primary Care R&D, Barts & the London, Queen Mary's School of Medicine and Dentistry, University of London

Ms Bec Hanley, Freelance Consumer Advocate, Hurstpierpoint, West Sussex

Professor Alan Horwich, Director of Clinical R&D, The Institute of Cancer Research, London

Dr Phillip Leech, Principal Medical Officer for Primary Care, Department of Health, London

Mr George Levvy, Chief Executive, Motor Neurone Disease Association, Northampton

Professor James Lindesay, Professor of Psychiatry for the Elderly, University of Leicester

Dr Mike McGovern, Senior Medical Officer, Heart Team, Department of Health, London

Dr John C Pounsford, Consultant Physician, North Bristol NHS Trust

Professor Mark Sculpher, Professor of Health Economics, Institute for Research in the Social Services, University of York

Dr L David Smith, Consultant Cardiologist, Royal Devon & Exeter Hospital

Professor Norman Waugh, Professor of Public Health, University of Aberdeen

Expert Advisory Network

Members

Mr Gordon Aylward,
Chief Executive,
Association of British Health-Care Industries, London

Ms Judith Brodie,
Head of Cancer Support Service, Cancer BACUP, London

Mr Shaun Brogan,
Chief Executive, Ridgeway Primary Care Group, Aylesbury, Bucks

Ms Tracy Bury,
Project Manager, World Confederation for Physical Therapy, London

Mr John A Cairns,
Professor of Health Economics, Health Economics Research Unit, University of Aberdeen

Professor Howard Stephen Cuckle,
Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds

Professor Nicky Cullum,
Director of Centre for Evidence Based Nursing, University of York

Dr Katherine Darton,
Information Unit, MIND – The Mental Health Charity, London

Professor Carol Dezateux,
Professor of Paediatric Epidemiology, London

Professor Martin Eccles,
Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne

Professor Pam Enderby,
Professor of Community Rehabilitation, Institute of General Practice and Primary Care, University of Sheffield

Mr Leonard R Fenwick,
Chief Executive, Newcastle upon Tyne Hospitals NHS Trust

Professor David Field,
Professor of Neonatal Medicine, Child Health, The Leicester Royal Infirmary NHS Trust

Mrs Gillian Fletcher,
Antenatal Teacher & Tutor and President, National Childbirth Trust, Henfield, West Sussex

Ms Grace Gibbs,
Deputy Chief Executive, Director for Nursing, Midwifery & Clinical Support Servs., West Middlesex University Hospital, Isleworth, Middlesex

Dr Neville Goodman,
Consultant Anaesthetist, Southmead Hospital, Bristol

Professor Robert E Hawkins,
CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester

Professor F D Richard Hobbs,
Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham

Professor Allen Hutchinson,
Director of Public Health & Deputy Dean of ScHARR, Department of Public Health, University of Sheffield

Professor Rajan Madhok,
Medical Director & Director of Public Health, Directorate of Clinical Strategy & Public Health, North & East Yorkshire & Northern Lincolnshire Health Authority, York

Professor David Mant,
Professor of General Practice, Department of Primary Care, University of Oxford

Professor Alexander Markham,
Director, Molecular Medicine Unit, St James's University Hospital, Leeds

Dr Chris McCall,
General Practitioner, The Hadleigh Practice, Castle Mullen, Dorset

Professor Alistair McGuire,
Professor of Health Economics, London School of Economics

Dr Peter Moore,
Freelance Science Writer, Ashtead, Surrey

Dr Andrew Mortimore,
Consultant in Public Health Medicine, Southampton City Primary Care Trust

Dr Sue Moss,
Associate Director, Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton, Surrey

Professor Jon Nicholl,
Director of Medical Care Research Unit, School of Health and Related Research, University of Sheffield

Mrs Julietta Patnick,
National Co-ordinator, NHS Cancer Screening Programmes, Sheffield

Professor Chris Price,
Visiting Chair – Oxford, Clinical Research, Bayer Diagnostics Europe, Cirencester

Ms Marianne Rigge,
Director, College of Health, London

Professor Sarah Stewart-Brown,
Director HSRU/Honorary Consultant in PH Medicine, Department of Public Health, University of Oxford

Professor Ala Szczepura,
Professor of Health Service Research, Centre for Health Services Studies, University of Warwick

Dr Ross Taylor,
Senior Lecturer, Department of General Practice and Primary Care, University of Aberdeen

Mrs Joan Webster,
Consumer member, HTA – Expert Advisory Network

Feedback

The HTA Programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278