KCE

Process notes

Search for Evidence & Critical Appraisal

Good Clinical Practice (GCP)

v. 2007-1

KCE Process notes

TABLE OF CONTENTS

1   INTRODUCTION

This document describes the methods of a literature review for the KCE. It provides guidance for reviewers on the various steps of the search, appraisal and presentation of the results.

New evidence may change some of the recommendations made, thereby researchers should consider this as a 'living document' for which yearly updates will be required. Next update is scheduled for June 2008.

This document is mainly based on four sources of information:

1. The Cochrane Collaboration Handbook (Higgins and Green 2006)
2. SIGN 50 (SIGN 2004)
3. Review Methods and Resources from CRD (Centre for Reviews and Dissemination (CRD) 2007)
4. The QUOROM statement (Moher et al. 1999)

2   METHODOLOGICAL APPROACH TO LITERATURE SEARCH

An evidence report consists of the following steps:

1. Formulating the problem and developing a protocol
2. Locating studies
3. Selecting studies
4. Quality assessment of studies
5. Collecting data, analysing and presenting results
6. Analysing and interpreting results

2.1   STEP 1 – FORMULATING THE QUESTION AND DEVELOPING A PROTOCOL

A protocol for carrying out a review is equivalent to, and as important as, a protocol for a primary research study. A review is less likely to be biased if the questions are well developed beforehand, and the methods that will be used to answer them are decided on before gathering the necessary data and drawing inferences. In the absence of a protocol, it is possible that study selection and analysis will be unduly driven by (a presumption of) the findings.

2.1.1   Objectives

The review should begin with a precise statement of the primary aim of the review, including the intervention(s) or test(s) reviewed and the targeted problem. This may be followed by a series of specific objectives relating to different participant groups, different comparisons of interventions or different outcome measures.

Definition of a set of clear and focused clinical questions is fundamental to the successful completion of a review. It is also important to be realistic about the number of questions that can be addressed in a single review if the final product is not to be too large to be useable. A large number of key questions also implies a very high workload for the reviewers, and care must be taken to ensure that this is kept within manageable limits.

2.1.2   Key components of a question

There are several key components to a well-formulated question (Richardson et al. 1995; Counsell 1997). A clearly defined question should specify

- the population type (participants),
- the types of interventions or exposures,
- the types of outcomes that are of interest.

A well known acronym used in this context is PICO (CEBM 2007). In addition, the types of studies that are relevant to answering the question should be specified.

An example is provided in appendix.

In general the more precise one is in defining components, the more focused the review. Equal precision in addressing each component is not necessary. For example, one may want to concentrate on various treatments for a particular stage of breast cancer, or alternately to focus on a particular drug for any stage of breast cancer. In the former example, the stage and severity of the disease would have to be defined very precisely within the 'Types of participants'. On the contrary, in the latter example, the treatment formulation would have to be defined very precisely within the 'Types of intervention'.

2.1.3   Methods outlined for the review

In the protocol, details on the methods of the review should be outlined. Essential components are the search terms that will be used, the databases that will be searched, and the selection criteria by which studies will be in or excluded from the review.

Subsequently, the methods for quality assessments should be described, as well as the consequences of the appraisal, e.g. will low quality studies be excluded from the review, treated separately or included with the good quality studies.

Finally, a description of the data extraction and possible analyses should be included.

2.2   STEP 2 – LOCATING STUDIES

A search strategy consists of several aspects. The research question should be used as a guide to direct the search strategy. For electronic searches, it is important to list the databases in which studies are sought, the terms used and filters applied and the dates on which the searches were performed to make it reproducible. Other sources can be consulted in order to identify all relevant studies. These include reference lists from relevant primary and review articles, journals, grey literature and conference proceedings, research registers, researchers and manufacturers and the internet.

In practice, it is rare for a single search to cover all the questions being addressed within a review. Different questions may be best answered by different databases, or may rely on different levels of evidence. Authors are encouraged to take an iterative approach to the search, carrying out a search for high level evidence in first instance. After the results of this search have been evaluated, the questions may be redefined and subsequent searches focused on the most appropriate sources and study types.

In some cases good quality, directly relevant evidence synthesis (secondary sources) such as good quality systematic reviews or Health Technology Assessments (HTA) will have been produced on some of the issues that fall within the remit of the review. In these circumstances reference will be made to the existing evidence rather than repeating work that has already been completed. All HTA reports or systematic reviews that are identified must be evaluated on their quality and be shown to have followed an acceptable methodology before they can be considered for use in this way.

In other cases existing evidence may not be directly relevant to the KCE, or may be found to have methodological weaknesses. In those cases, existing evidence can not be used in the review. But, excluded systematic reviews or HTA reports are a useful source of references that might be used later on in the review.

As a result, literature searches for the KCE should follow an iterative approach, searching for evidence synthesis first and subsequently complementing this by searching for original studies. The various sources are listed in the following paragraph.

2.2.1   Electronic searches

The three electronic bibliographic databases generally considered as the richest sources of primary studies - MEDLINE, EMBASE, and CENTRAL - are essential in any literature review for the KCE. Multiple other electronic bibliographic databases exist.

Systematic reviews can be found in the Cochrane Database for Systematic Reviews, in DARE or in Medline. Search strategies have been developed to enhance identification of these types of publication (Boynton et al. 1998).

HTA reports can be found in the HTA database of CRD, or at individual agencies' sites (see INAHTA's website under members[a]).

Specifically for drugs and technology reviews, data from the U.S. Food and Drug Administration[b] or the European Medicines Agency[c] can be helpful.

Details of other databases that might contain eligible records are listed at the website of NICE (NICE 2007).

The selection of which source to use depends on the research question. In addition, providing an exhaustive list of all possible sources is not possible. Authors may want to consult the CRD website, where a large number of electronic or other databases are listed, with a description of their scope (Centre for Reviews and Dissemination (CRD)).

2.2.1.1   Core list of resources

- **MEDLINE:** Records from 5000 journals (37 languages) in the of biomedical field, from 1966 onwards (Old Medline, included in PubMed, from 1950).
- **EMBASE:** Records from 5000 journals (70 countries) in biomedical field, from 1974 onwards.
- **CENTRAL** (The Cochrane Controlled Trials Register): Records of randomised controlled trials and controlled clinical trials in healthcare identified through the work of the Cochrane Collaboration including large numbers of records from MEDLINE and EMBASE as well as much material not covered by these databases (Dickersin et al. 2002).

2.2.1.2   Specific resources

- **Centre for Reviews and Dissemination (CRD**) includes the Database of Abstract of Reviews of Effects (DARE), the NHS Economic Evaluation Database (NHS EED) and the Health Technology Assessment Database (HTA). The 3 different CRD databases include structured abstracts, identified by regular searching of bibliographic databases, and hand searching of key journals.
- **Clinical Trial Registries:** several initiatives have been taken recently to register ongoing trials. The WHO Registry Platform is a project within the World Health Organization, to unite all possible trial register (WHO 2007).

  Other examples are ClinicalTrials.gov[d], TrialsCentral[TMe], Current Controlled Trials (www.controlled-trials.com). or Eudract[f]. Ongoing trials may have limited use as a means of identifying studies relevant to systematic reviews, but may be important so that when a review is later updated, these studies can be assessed for possible inclusion.

---

[a] http://www.inahta.org
[b] http://www.fda.gov/
[c] http://www.emea.europa.eu/
[d] http://clinicaltrials.gov/
[e] http://www.trialscentral.org
[f] http://eudract.emea.europa.eu/

- **CINAHL** (Cumulative Index to Nursing and Allied Health Literature): Records of literature on all aspects of nursing and allied health disciplines.

- **PEDro** (Physiotherapy Evidence Database): Records of RCTs, systematic reviews and evidence-based clinical practice guidelines in physiotherapy, from 1929. Most trials on the database have been rated for quality to quickly discriminate between trials which are likely to be valid and interpretable and those which are not.

- **PsycInfo:** Records of research in psychology and related behavioural and social sciences, from 1967.

- **Clinical Practice Guidelines:** Best known is the National Guideline Clearinghouse, a USA based database of primarily clinical practice guidelines in the English language. Other websites exist as well, such as the website of the Guideline International Network (GIN), but there is not one single database or search engine covering all possible CPGs. Several guidelines are never published in peer-reviewed journals, rendering the search for CPGs often cumbersome. Often, specific guidelines can only be retrieved through local websites of scientific associations or government agencies (list available on the CEBAM website).

2.2.2   Search terms

Constructing an effective combination of search terms for searching electronic databases requires a structured approach. One such approach involves breaking down the review question into 'facets', such as population, interventions, outcomes and study designs.

The next stage is to identify the search terms in each 'facet' which best capture the subject. The group of search terms covering each facet of the review question should include a range of textwords (free text) in the title or abstract of studies as well as any available subject indexing terms that are assigned by the database producer (e.g. MeSH).

- Text words and their variants can be identified from reading relevant reviews and primary studies identified during earlier searches.

- Information on the subject indexing used by databases can be found by consulting the relevant indexing manuals and by noting the manner in which key retrieved articles have been indexed by a given database.

The final search strategies will be developed by an iterative process in which groups of terms are used, perhaps in several permutations, to identify the combination of terms that seems most sensitive in identifying relevant studies. This requires skilled adaptation of search strategies based on knowledge of the subject area, the subject headings and the combination of 'facets' which best capture the topic.

An example:

The question: In patients undergoing hip replacement, to what extent is the risk of post-operative infection reduced by antimicrobial prophylaxis?

Break down of the question into 'facets'

• Population	Patients undergoing hip replacement
• Interventions	Antimicrobial prophylaxis
• Outcome	Post-operative infection
• Study design	Randomised Controlled Trials

Subject headings, synonyms or spelling variants for *post-operative infection:*

Text words	Medical Subject Headings (MeSH)
Bacterial infection	BACTERIAL INFECTIONS
Postoperative complication(s)	POSTOPERATIVE COMPLICATIONS
Wound infection	SURGICAL WOUND INFECTION
	PROSTHESIS-RELATED INFECTIONS
Septicaemia	SEPSIS
Bacterial contamination	INFECTION CONTROL

2.2.3   Filters

In systematic reviews, if time and resources allow, specificity is often sacrificed in favour of sensitivity, to maximize the yield of relevant articles. Therefore, it is not unusual to retrieve large numbers (possibly thousands) of bibliographic references for consideration for inclusion in an extensive systematic review. This means that reviewers may have to spend a lot of time scanning references to identify perhaps a limited number of relevant studies.

Search filters are available to focus the search according to the type of study that is sought, for example to focus on randomized controlled trials, on diagnostic accuracy studies, on prognostic studies or on systematic reviews (see example in Appendix).

Source of filters:

- PubMed (NLM 2007)
- InterTASC (InterTASC Information Specialists' Sub-Group)
- SIGN (SIGN 2007)
- HiRU (Health Information Research Unit 2007)
- OVID or Embase.com

2.2.4   Checking reference lists

Authors should check the reference lists of articles obtained (including those from previously published systematic reviews) to identify relevant reports. The process of following up references from one article to another is generally an efficient means of identifying studies for possible inclusion in a review. Because investigators may selectively cite studies with positive results (Gotzsche 1987; Ravnskov 1992), reference lists should never be used as a sole approach to identifying reports for a review, but rather as an adjunct to other approaches.

2.2.5   Evidence on adverse effects

The first sources to investigate for information on adverse effects are reports from trials or other studies included in the systematic review. Excluded reports might also provide some useful information.

There are a number of specific sources of information on adverse effects of drugs, including

- Current Problems produced by the UK Medicines Control Agency[g],
- MedWatch produced by the US Food and Drug Administration, (FDA)
- Australian Adverse Drug Reactions Bulletin (Adverse Drug Reactions Advisory Committee).

In Belgium, there is currently no public database on adverse drug events. Other regulatory authorities (such as the websites of FDA and EMEA) and the drug manufacturer may also be able to provide some information. Information on adverse effects might also be sought from other types of studies than those considered appropriate for the systematic review (e.g.

---

[g] http://www.open.gov.uk/mca

cohort and case-control studies, uncontrolled trials, case series and case reports). However, all such studies and reports are subject to bias to a greater extent than randomized trials, and findings must be interpreted with caution.

2.2.6   Documenting a search strategy

The search strategy for electronic databases should be described in sufficient detail in a review that the process could be replicated.

The bibliographic databases searched, the dates and periods searched and any constraints, such as language should be stated. The full search strategies for each database should be listed in an additional table or in the appendix.

The template required by KCE is provided in appendix.

2.3   STEP 3 – SELECTING STUDIES

Study selection is a multi-stage process. The process by which studies will be selected for inclusion in a review should be described in the review protocol.

2.3.1   Evidence sifting

Before any papers are acquired for evaluation, sifting of the search output is carried out to eliminate irrelevant material.

- Papers that are clearly not relevant to the key questions are eliminated based on their title.
- Abstracts of remaining papers are then examined and any that are clearly not appropriate study designs, or that fail to meet specific methodological criteria, will be also eliminated at this stage.
- All reports of studies that are identified as potentially eligible must then be assessed in full text to see whether they meet the inclusion criteria for the review.

The reproducibility of this process should be tested in the initial stages of the review, and if reproducibility is shown to be poor more explicit criteria may have to be developed to improve it.

Authors must decide whether more than one author will assess the relevance of each report. Whatever the case, the number of people assessing the relevance of each report should be stated in the Methods section of the review. Some authors may decide that assessments of relevance should be made by people who are blind or masked to the journal from which the article comes, the authors, the institution, and the magnitude and direction of the results by editing copies of the articles (Berlin 1997; Berlin, Miles, and Crigliano 1997). However, this takes much time, and may not be warranted given the resources required and the uncertain benefit in terms of protecting against bias (Berlin 1997).

2.3.2   Inclusion and exclusion

The final inclusion/exclusion decisions should be made after retrieving the full texts of all potentially relevant citations. Reviewers should assess the information contained in these reports to see whether the criteria have been met or not. Many of the citations initially included may be excluded at this stage.

The criteria used to select studies for inclusion in the review must be clearly stated:

TYPES OF STUDIES

Eligible study designs should be stated here, along with any thresholds for inclusion based on the conduct or quality of the studies. For example, *'All randomised controlled comparisons'* or *'All randomised controlled trials with blind assessment of outcome'*. Exclusion of particular types of randomised studies (for example, cross-over trials) should be justified.

It is generally for authors to decide which study design(s) to include in their review. Some reviews are more restrictive, and include only randomized trials, while others are less restrictive, and include other study designs as well, particularly when few randomized trials addressing the topic of the review are identified. For example, many of the reviews from the

Cochrane Effective Practice and Organization of Care (EPOC) Collaborative Review Group include before-and-after studies and interrupted time series in addition to randomized and quasi-randomized trials.

TYPES OF PARTICIPANTS

The diseases or conditions of interest should be described here, including any restrictions on diagnoses, age groups and settings. Subgroup analyses should not be listed here.

TYPES OF INTERVENTIONS

Experimental and control interventions should be defined here, making it clear which comparisons are of interest. Restrictions on dose, frequency, intensity or duration should be stated. Subgroup analyses should not be listed here.

TYPES OF OUTCOME MEASURES

Note that outcome measures do not always form part of the criteria for including studies in a review. If they do not, then this should be made clear. Outcome measures of interest should be listed in this section whether or not they form part of the inclusion criteria.

For most reviews it will be worthwhile to pilot test the inclusion criteria on a sample of articles (say ten to twelve papers, including ones that are thought to be definitely eligible, definitely not eligible and questionable). The pilot test can be used to refine and clarify the inclusion criteria, train the people who will be applying them and ensure that the criteria can be applied consistently by more than one person.

Even when explicit inclusion criteria have been specified, decisions concerning the inclusion of individual studies remain relatively subjective. There is evidence that using at least two authors has an important effect on reducing the possibility that relevant reports will be discarded (Edwards et al. 2002). Agreement between assessors may be formally assessed mathematically using Cohen's Kappa (a measure of chance-corrected agreement). Many disagreements may be simple oversights, whilst others may be matters of interpretation. These disagreements should be discussed, and where possible resolved by consensus after referring to the protocol. If disagreement is due to lack of information, the authors may have to be contacted for clarification. Any disagreements and their resolution should be recorded. The influence of uncertainty about study selection may be investigated in a sensitivity analysis.

It is useful to construct a list of excluded studies at this point, detailing the reason for each exclusion. This list may be included in the report of the review as an appendix. The final report of the review should also include a flow chart or a table detailing the studies included and excluded from the review. In appendix a flow chart is provided for documenting study selection. If resources and time allow, the lists of included and excluded studies may be discussed with the expert panel. It may be useful to have a mixture of subject experts and methodological experts assessing inclusion.

2.4   STEP 4 – CRITICAL APPRAISAL OF THE EVIDENCE

Each report, article or guideline that is selected as a potential source of evidence is critically appraised based on the following questions:

1. Is the article **relevant** to the subject?

2. Are the article's results **valid**?

3. Are the article's results **important** for answering the question?

The methodological assessment is based on a number of key questions that focus on those aspects of the study design that have a significant influence on the validity of the results reported and conclusions drawn. These key questions differ between study types, and a range of checklists can be used to bring a degree of consistency to the assessment process:

- For HTA reports, the INAHTA checklist is recommended (see appendix). Key questions are the adequacy of the literature search and quality appraisal of the selected studies.

- For systematic reviews, the checklist of the Dutch Cochrane Centre can be used (see appendix). These checklists were translated into a French version by experts of the KCE (available on demand). Key aspects are similar to those for HTA reports (adequacy of literature search and quality appraisal).

- Examples of checklists for primary studies, such as those from Dutch Cochrane and SIGN, can be found in appendix. For randomised controlled trials, the randomisation process, blinding of the outcome assessors and an intention-to-treat-analysis are important quality criteria. For observational studies, blinded assessment of the outcomes and adequate dealing with confounders are essential.

- For the critical appraisal of clinical practice guidelines, the AGREE instrument should be used (see appendix). This checklist consists of 23 items, divided into 6 domains. The instrument does not assess the clinical content of the recommendations. The instructions in the introduction of the instrument should be read carefully before starting the appraisal. Each guideline should be appraised by at least two appraisers. The AGREE instrument does not provide thresholds for acceptable or unacceptable guidelines based on quality. In general, guidelines are to be considered as indirect sources of evidence. Unless they are based on a good quality systematic review of the literature, guidelines are prone to be authority-biased and to represent expert or stakeholders opinions. If a systematic review underpinning the statements of the guideline is available, the review should be used to appraise the quality of the evidence, using the systematic review checklist for quality appraisal.

The critical appraisal process inevitably involves a degree of subjective judgement. To minimise any potential bias resulting from this, it is recommended that each study is evaluated independently by two members of the project group. Any differences in assessment should be discussed. Where differences cannot be resolved, an independent reviewer or an experienced member of the staff will arbitrate to reach an agreed quality assessment. Validation by a third researcher experienced in literature review is highly recommended as part of the quality control process.

2.5   STEP 5 – DATA EXTRACTION, TABLES OF EVIDENCE, AND LEVEL OF EVIDENCE

Data extraction implies the process of extracting the information from the selected studies that will be ultimately reported. In order to allow an efficient data extraction, the process should be detailed in the protocol before the literature search is started. Key components of the data extraction include:

- information about study reference(s) and author(s);

- verification of study eligibility;

- study characteristics:
  - study methods
  - participants
  - interventions
  - outcomes measures and results

All validated studies identified from the systematic literature review relating to each key search question are summarized into evidence tables. The content of the evidence tables is determined by the entire project group. Completion for all retained articles is done by one member of the project group.

As a basis for the tables of evidence, the Summary of Findings (SoF) tables of the Cochrane Collaboration are used (see appendix). SoF tables have been suggested to help readers quickly focus on the key results and access information that is needed to inform a decision. The SoF table includes information on each of the main outcomes addressed in the guideline. The number of patients and trials, the control group risk, the effect size (relative and absolute), and the quality of the evidence are presented for each main outcome separately.

To allocate a level of evidence, the GRADE system is used (see appendix). The quality of the evidence ranges from high, over moderate and low, to very low. The study design is the major determinant for the level of evidence, but this level can be lowered or increased depending on the quality of the study.

2.6   STEP 6 – ANALYSING AND INTERPRETING RESULTS

Once the eligible studies are selected and quality appraised, the magnitude of the intervention effect should be estimated. The best way to do this is by performing a meta-analysis (i.e. the statistical combination of results from two or more separate studies), although this is not always feasible. An interesting tool for doing a limited meta-analysis is the free Review Manager software of the Cochrane Collaboration[h].

The starting point of the analysis and interpretation of the study results involves the identification of the data type for the outcome measurements. Five different types of outcome data can be considered:

- dichotomous data: two possible categorical response;
- continuous data
- ordinal data: several ordered categories;
- counts and rates calculated from counting the numbers of events that each individual experiences;
- time-to-event data

Only dichotomous data will be addressed here. Dichotomous outcome data arise when the outcome for every study participant is one of two possibilities, for example, dead or alive. These data can be summarised in a 2x2 table:

	Outcome		
	Yes	No	
Intervention	a	b	a + b
Control	c	d	c + d
	a + c	b + d	

The most commonly encountered effect measures used in clinical trials with dichotomous data are:

- Relative risk (RR): the ratio of the risk (i.e. the probability with which the outcome will occur) of the outcome in the two groups, or [a/(a+b)]/[c/(c+d)]. For example, a RR of 3 implies that the outcome with treatment is three times more likely to occur than without treatment;
- Absolute risk reduction (ARR): the absolute difference of the risk of the outcome in the two groups, or [a/(a+b)]-[c/(c+d)];
- Number needed to treat (NNT): the number of persons that need to be treated with the intervention in order to prevent one additional outcome, or 1/ARR.
- For diagnostic accuracy studies, the results will be expressed as
- Sensitivity: the proportion of true positives correctly identified by the test: Sens=a/a+c
- Specificity: the proportion of true negatives correctly identified by the test: Spec=d/b+d
- Positive predictive value: the proportion of patients with a positive test result correctly diagnosed: PPV=a/a+b
- Negative predictive value: the proportion of patients with a negative test result correctly diagnosed: NPV=d/c+d
- Likelihood ratio: likelihood that a given test result would be expected in a patient with the target disorder compared to the likelihood that that same result would be expected in a patient without the target disorder LR+=(a/a+c)/(b/b+d); LR-=(c/a+c)/(d/b+d)

---

[h] http://www.cc-ims.net/RevMan

- Diagnostic odds ratio: ratio of the odds of having a positive index test result in a patient with the target condition over the odds of having this test result in a patient without the target condition: OR=ad/bc

	Target condition Positive	Target condition Negative
Index test positive	a	b
Index test negative	c	d

As discussed above, other types than dichotomous data are possible, each with their own outcome measures and statistics. It is beyond the scope of this document to describe and discuss all these types. Interested readers are referred to textbooks such as Practical statistics for medical research (Altman 1991) Modern Epidemiology (Rothman and Greenland 1998) and Clinical epidemiology : a basic science for clinical medicine (Sackett 1991) .

3   REPORTING OF A LITERATURE SEARCH

A literature search should be reproducible and therefore explicitly documented. The report of a literature search should contain the following items:

1. Description of the search methodology:

    a. Search protocol

        i.   Search question

        ii.   Searched databases

        iii.   Search terms, their combinations and the restrictions used (e.g. language, date)

        iv.   In- and exclusion criteria for the selection of the studies

    b. Quality appraisal methodology

    c. Data extraction methodology

2. Description of the search results:

    a. Number of retrieved articles, in- and excluded studies, and reasons for exclusion; use of flow chart

    b. Results of quality appraisal

    c. Evidence tables for each search question

4   REFERENCES

Adverse Drug Reactions Advisory Committee. *Australian Adverse Drug Reactions Bulletin* [Home Page on the Internet]. Therapeutic Goods Administration, aug. 01 2007 [cited 2007-09-20]. Available from http://www.tga.gov.au/adr/aadrb.htm.

AGREE Collaboration. 2001. AGREE Instrument: Appraisal of Guidelines for research & Evaluation: Agree collaboration.

Altman, Douglas G. 1991. *Practical statistics for medical research*. 1st ed. London ; New York: Chapman and Hall.

Berlin, J. A. 1997. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania Meta-analysis Blinding Study Group. Review of Lancet. 1997 Sep 20;350(9081):892. PMID: 9310638. *Lancet* 350 (9072):185-6.

Berlin, JA, CG Miles, and MD Crigliano. 1997. Does blinding of readers affect the results of meta-analyses? Results of a randomized trial. *Online J Curr Clin Trials*.

Boynton, Janette, Julie Glanville, David McDaid, and Carol Lefebvre. 1998. Identifying systematic reviews in MEDLINE: developing an objective approach to search strategy design. *Journal of Information Science* 24:137-154.

CEBM. 2007. *Asking Focused Clinical Questions* [HomePage on the Internet]. Center for Evidence-Based Medicine 2007 [cited 2007-08-27 2007]. Available from http://www.cebm.net/index.aspx?o=1036.

Centre for Reviews and Dissemination (CRD). *Finding studies for systematic reviews: a checklist for researchers* [Home Page on the Internet]. University of York, oct. 2006 [cited 2007-09-20]. Available from http://www.york.ac.uk/inst/crd/revs.htm.

————. *Review Methods and Resources*. University of York,, 2007-08-23 2007 [cited 2007-08-27. Available from http://www.york.ac.uk/inst/crd/crdreview.htm.

Counsell, C. 1997. Formulating questions and locating primary studies for inclusion in systematic reviews. *Ann Intern Med* 127 (5):380-7.

Dickersin, Kay, Eric Manheimer, Susan Wieland, Karen A. Robinson, Carol Lefebvre, and Steve McDonald. 2002. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* 25 (1):38-64.

Edwards, Phil, Mike Clarke, Carolyn DiGuiseppi, Sarah Pratap, Ian Roberts, and Reinhard Wentz. 2002. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. *Stat Med* 21 (11):1635-40.

FDA. *MedWatch: The FDA Safety Information and Adverse event Reporting Program* [Home Page on the Internet]. U.S. Food and Drugs Administration (FDA), sept. 18 2007 [cited 2007-09-20]. Available from http://www.fda.gov/medwatch/.

Gotzsche, P. C. 1987. Reference bias in reports of drug trials. *Br Med J (Clin Res Ed)* 295 (6599):654-6.

Guyatt, Gordon, David Gutterman, Michael H. Baumann, Doreen Addrizzo-Harris, Elaine M. Hylek, Barbara Phillips, Gary Raskob, Sandra Zelman Lewis, and Holger Schunemann. 2006. Grading strength of recommendations and quality of evidence in clinical guidelines: report from an american college of chest physicians task force. Review of Chest. 2006 Jan;129(1):10-2. PMID: 16424406

Chest. 2006 Jan;129(1):7-10. PMID: 16424405. *Chest* 129 (1):174-81.

Health Information Research Unit. 2007. *Search Strategies for MEDLINE in Ovid Syntax and the PubMed translation* [Home Page on the Internet]. McMaster University 2007 [cited 2007-08-27 2007]. Available from http://hiru.mcmaster.ca/hedges/index.htm.

Higgins, Julian PT, and Sally Green. 2006. *Cochrane Handbook for Systematic Reviews of Interventions*. 4.2.6 ed. Vol. Issue 4, *The Cochrane Library*. Chichester, UK: John Wiley & Sons.

INAHTA Secretariat. 2001. A checklist for health technology assessment reports.

InterTASC Information Specialists' Sub-Group. 2007. *Search Filter Resource* [Home Page on the Internet]. University of York, aug. 07 2007 [cited 2007-08-27 2007]. Available from http://www.york.ac.uk/inst/crd/intertasc/sr.htm.

Moher, D., D. J. Cook, S. Eastwood, I. Olkin, D. Rennie, and D. F. Stroup. 1999. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. Review of Lancet. 1999 Nov 27;354(9193):1834-5. PMID: 10584713

NICE. *Databases* [Home Page on the Internet]. National Institute for Health and Clinical Excellence, 2007 [cited 2007-09-20]. Available from http://www.nice.org.uk/page.aspx?o=516408.

NLM. 2007. *Special Queries: Directory of Topic-Specific PubMed Queries* [Home Page on the Interne]. National Library of Medicine (NLM), may 23 2007 [cited 2007-09-20 2007]. Available from http://www.nlm.nih.gov/bsd/special_queries.html.

Oxman, Andy , Julian Higgins, and Paul Glasziou. 2006. Summary of Findings and plain language summaries.

Ravnskov, U. 1992. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. Review of BMJ. 1992 Aug 15;305(6850):420-1; author reply 422. PMID: 1392931

Richardson, W. S., M. C. Wilson, J. Nishikawa, and R. S. Hayward. 1995. The well-built clinical question: a key to evidence-based decisions. *ACP J Club* 123 (3):A12-3.

Rothman, Kenneth J., and Sander Greenland. 1998. *Modern epidemiology*. 2nd ed. Philadelphia: Lippincott-Raven.

Sackett, David L. 1991. *Clinical epidemiology : a basic science for clinical medicine*. 2nd ed. Boston: Little, Brown.

SIGN. 2004. *SIGN 50: A guideline developers' handbook*. Edinburgh: Scottish Intercollegiate Guidelines Network (SIGN).

———. *Search filters* [Home Page on the Internet]. Scottish Intercollegiate Guidelines Network (SIGN) 2007 [cited 2007-09-20. Available from http://www.sign.ac.uk/methodology/filters.html.

Whiting, Penny, Anne W. S. Rutjes, Johannes B. Reitsma, Patrick M. M. Bossuyt, and Jos Kleijnen. 2003. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 3:25.

WHO. *International Clinical Trials Registry Platform (ICTRP)* [Home Page on the Internet]. World Health Organization 2007 [cited 2007-09-20]. Available from http://www.who.int/ictrp/en/.

5 APPENDICES

APPENDIX 1: EXAMPLE OF PICO

Choice β blocker

A patient diagnosed with hypertension and admitted for a laparoscopic knee surgery, asks you whether Tenormin, the antihypertensive drug he is taking, is better than Selozok, the one his neighbour is taking.

Patient population	patient with hypertension
Intervention	atenolol
Comparison:	metoprolol
Outcomes	(cardiovascular) mortality

APPENDIX 2: EXAMPLE FILTER FOR SYSTEMATIC REVIEWS IN MEDLINE (OVID)

The following search terms can be added to the topic-related search terms, in order to identify systematic reviews only. Applying a filter reduces the number of articles that has to be read, in order for one article to be included in the review.

1.	Meta-Analysis/
2.	meta analy$.tw.
3.	metaanaly$.tw.
4.	meta analysis.pt.
5.	(systematic adj (review$1 or overview$1)).tw.
6.	exp Review Literature/
7.	or/1-6
8.	cochrane.ab.
9.	embase.ab.
10.	(psychlit or psyclit).ab.
11.	(psychinfo or psycinfo).ab.
12.	(cinahl or cinhal).ab.
13.	science citation index.ab.
14.	bids.ab.
15.	cancerlit.ab.
16.	or/8-15
17.	reference list$.ab.
18.	bibliograph$.ab.
19.	hand-search$.ab.
20.	relevant journals.ab.
21.	manual search$.ab.
22.	or/17-21
23.	selection criteria.ab.
24.	data extraction.ab.
25.	23 or 24
26.	review.pt.
27.	25 and 26
28.	comment.pt.
29.	letter.pt.
30.	editorial.pt.
31.	animal/
32.	human/
33.	31 not (31 and 32)
34.	or/28-30,33
35.	7 or 16 or 22 or 27
36.	35 not 34

APPENDIX 3 DOCUMENTING A SEARCH STRATEGY

Author	
Name	
Project number	
Project name	
Keywords	

Date (day month year)	
Database (name +provider ; eg Medline OVID)	
Search Strategy (attention, for PubMed, check « Details »)	
Note	

The second table must be copied as many times as necessary.

APPENDIX 4 FLOW DIAGRAM OF STUDY SELECTION PROCESS

From QUOROM statement (Moher et al. 1999)

```
┌─────────────────────────────────────┐
│  Potentially relevant studies        │
│  identified and screened for         │
│  retrieval (n=…)                     │
└─────────────────────────────────────┘
              │
              │      ┌─────────────────────────────────┐
              ├─────▶│  Studies excluded with reason x │
              │      │  (n=…)                          │
              │      │        with reason y (n=  )     │
              ▼      └─────────────────────────────────┘
┌─────────────────────────────────────┐
│  Studies retrieved for more detailed │
│  information (n=…)                   │
└─────────────────────────────────────┘
              │
              │      ┌─────────────────────────────────┐
              ├─────▶│  Studies excluded with reason x │
              │      │  (n=…)                          │
              │      │        with reason y (n=  )     │
              ▼      └─────────────────────────────────┘
┌─────────────────────────────────────┐
│  Potentially appropriate studies to  │
│  be included in the review (n=…)     │
└─────────────────────────────────────┘
              │
              │      ┌─────────────────────────────────┐
              ├─────▶│  Studies excluded with reason x │
              │      │  (n=…)                          │
              │      │        with reason y (n=  )     │
              ▼      └─────────────────────────────────┘
┌─────────────────────────────────────┐
│  Studies ultimately included in the  │
│  review (n=…)                        │
└─────────────────────────────────────┘
```
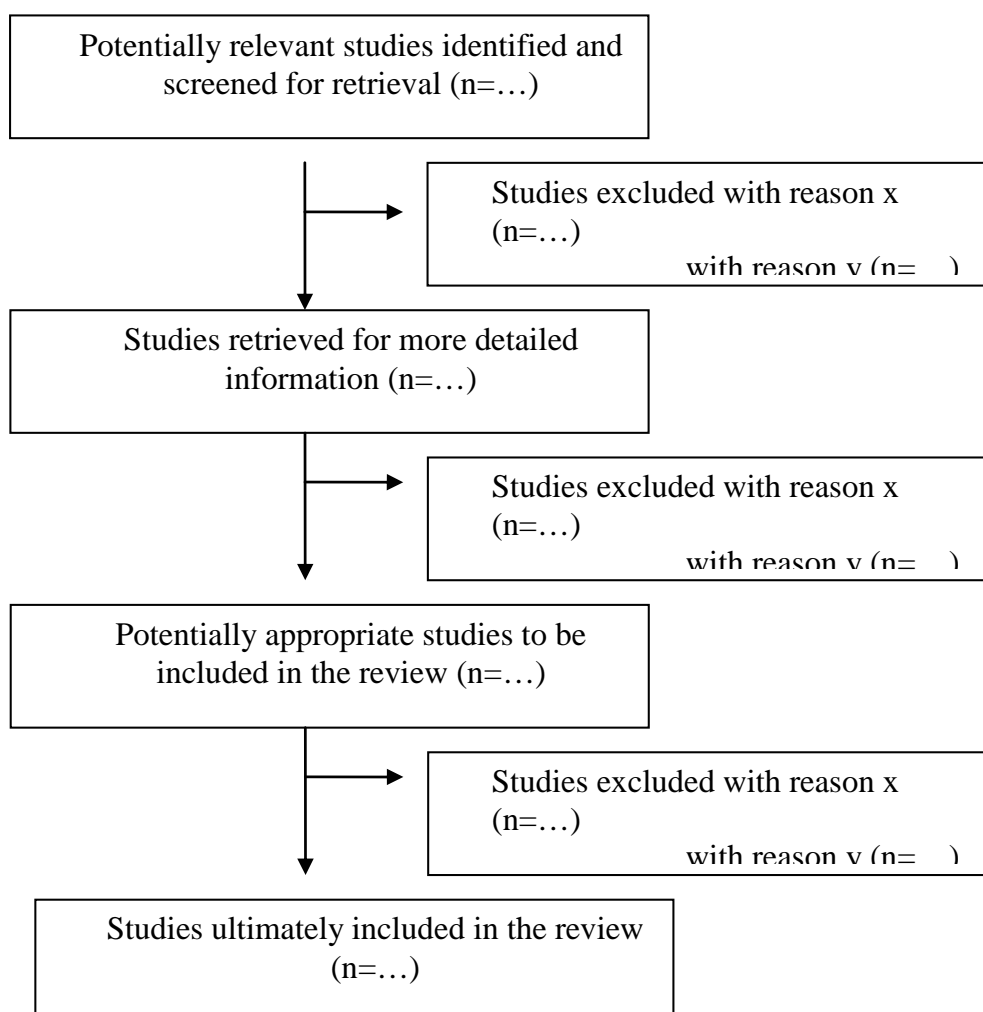
APPENDIX 5: QUALITY APPRAISAL CHECKLISTS

INAHTA CHECKLIST FOR THE APPRAISAL OF HTA REPORTS

From INAHTA Secretariat (INAHTA Secretariat 2001)

A summary for HTA reports

This summary form is intended as an aid for those who wish to make a record of the extent to which a health technology assessment report meets the 17 questions given in the checklist.

It is NOT intended as a scorecard to rate the standard of HTA reports — reports may be valid and useful without meeting all the criteria that have been listed.

Item	Yes	Partly	No
Preliminary			
1. Appropriate contact details for further information?			
2. Authors identified?			
3. Statement regarding conflict of interest?			
4. Statement on whether report externally reviewed?			
5. Short summary in non-technical language?			
Why?			
6. Reference to the question that is addressed and context of the assessment?			
7. Scope of the assessment specified?			
8. Description of the health technology?			
How?			
9. Details on sources of information?			
10. Information on selection of material for assessment ?			
11. Information on basis for interpretation of selected data?			
What?			
12. Results of assessment clearly presented?			
13. Interpretation of the assessment results included? What then?			
14. Findings of the assessment discussed?			
15. Medico-legal implications considered?			
16. Conclusions from assessment clearly stated?			
17. Suggestions			

QUADAS CHECKLIST FOR DIAGNOSTIC ACCURACY STUDIES

From Whiting (Whiting et al. 2003)

Item	Yes	No	Unclear
1. Was the spectrum of patients representative of the patients who will receive the test in practice?			
2. Were selection criteria clearly described?			
3. Is the reference standard likely to correctly classify the target condition?			
4. Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?			
5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis?			
6. Did patients receive the same reference standard regardless of the index test result?			
7. Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?			
8. Was the execution of the index test described in sufficient detail to permit replication of the test?			
9. Was the execution of the reference standard described in sufficient detail to permit its replication?			
10 Were the index test results interpreted without knowledge of the results of the reference standard?			
11. Were the reference standard results interpreted without knowledge of the results of the index test?			
12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?			
13. Were uninterpretable/ intermediate test results reported?			
14. Were withdrawals from the study explained?			

DUTCH COCHRANE CHECKLISTS

- Dutch Cochrane Checklists Werkgroep. Beoordeling randomised controlled trial (RCT). Amsterdam: Dutch Cochrane Centre; 2002. Formulier II Available from: http://145.117.214.42/DCC/Downloads/RCT.pdf
- Dutch Cochrane Checklists Werkgroep. Beoordeling cohortonderzoek. Amsterdam: Dutch Cochrane Centre; 2002. Formulier III Available from: http://145.117.214.42/DCC/Downloads/cohort.pdf
- Dutch Cochrane Checklists Werkgroep. Beoordeling systematische review van RCT's. Amsterdam: Dutch Cochrane Centre; 2002. Formulier Va Available from: http://145.117.214.42/DCC/Downloads/SR-RCT.pdf

A French translation is available at KCE.

AGREE INSTRUMENT

From AGREE (AGREE Collaboration 2001)

- AGREE Collaboration. AGREE Instrument: Appraisal of Guidelines for research & Evaluation. Agree collaboration; 2001. (1 8981 8321 X) Available from: http://www.agreecollaboration.org/pdf/agreeinstrumentfinal.pdf

SIGN NOTES AND CHECKLISTS

- Methodology Checklist 1: Systematic Reviews and Meta-analyses
- Methodology Checklist 2: Randomised Controlled Trials
- Methodology Checklist 3: Cohort Studies
- Methodology Checklist 4: Case-control Studies
- Methodology Checklist 5: Diagnostic Studies
- Methodology Checklist 6: Economic Evaluations

http://www.sign.ac.uk/methodology/checklists.html

APPENDIX 6: SUMMARY OF FINDINGS TABLES

From OXMAN (Oxman, Higgins, and Glasziou 2006)

LINK:

http://www.cochrane.org/ccsg/SummaryofFindingstablesandplainlanguagesummaries.doc

EXAMPLE:

Question: Should antibiotics be used for acute otitis media in children?
Patient or population: Children without tympanostomy tubes, suffering from acute otitis media

Settings: The included trials were conducted in Europe and North America

Summary of findings

Outcome	No of Participants (No of trials)	Control group risk (Range)	Relative effect (95% CI)	Absolute effect	Quality	Comments
Pain after 1 day	717 (3)	38.5% (28 to 48%)	RR 1.02 (0.85 to 1.22)	Nil fewer/1 000	⊕⊕⊕⊕ High	
Pain between 2 & 7 days	2287 (9)	22.2% (8 to 72%)	RR 0.70 (0.60 to 0.81)	70 fewer/1 000	⊕⊕⊕⊕ High	Probably greater effect if fever or vomiting.
Mastoditis	2287 (9)	0%	-	-	⊕⊕⊕○ Moderate	Only one case of mastoditis was reported among 2287 patients in 9 trials (in an antibiotic treated group).
Glue ear at 3 months	370 (2)	26.1% (23 to 28%)	RR 0.80 (0.55 to 1.16)	-	⊕⊕⊕○ Moderate	Tympanometry in 2 trials only.
Adverse effects	938 (4)	10.5% (1 to 30%)	RR 1.60 (1.19 to 2.16)	62 more/1 000	⊕⊕⊕○ Moderate	Mostly diarrhoea, vomiting or rash.

APPENDIX 7: GRADE SYSTEM

From Guyatt (Guyatt et al. 2006)

Grade of Recommendation/ Description	Benefit vs. Risk and Burdens	Methodological Quality of Supporting Evidence	Implications
1A/ Strong recommendation, high quality evidence	Benefits clearly outweigh risk and burdens, or vice versa	RCTs without important limitations or overwhelming evidence from observational studies	Strong recommendation, can apply to most patients in most circumstances without reservation
1B/ Strong recommendation, moderate quality evidence	Benefits clearly outweigh risk and burdens, or vice versa	RCTs with important limitations (inconsistent results, methodological flaws, indirect, or imprecise) or exceptionally strong evidence from observational studies	Strong recommendation, can apply to most patients in most circumstances without reservation
1C/ Strong recommendation, low quality evidence	Benefits clearly outweigh risk and burdens, or vice versa	Observational studies or case series	Strong recommendation, but may change when higher quality evidence becomes available
2A/ Weak recommendation, high quality evidence	Benefits closely balanced with risks and burden	RCTs without important limitations or overwhelming evidence from observational studies	Weak recommendation, best action may differ depending on circumstances or patients' or societal values
2B/ Weak recommendation, moderate quality evidence	Benefits closely balanced with risks and burden	RCTs with important limitations (inconsistent results, methodological flaws, indirect, or imprecise) or exceptionally strong evidence from observational studies	Weak recommendation, best action may differ depending on circumstances or patients' or societal values
2C/ Weak recommendation, low quality evidence	Benefits closely balanced with risks and burden	Observational studies or case series	Very weak recommendation, other alternatives may be equally reasonable

Link: http://www.chestjournal.org/cgi/reprint/129/1/174?ijkey=1eeff3fc9de62d407c46a389d2e968d93532eceb

5

APPENDIX 8: USEFUL LINKS

- Cochrane: www.cochrane.org
  - o   Dutch Cochrane Centre: www.cochrane.nl
  - o   CEBAM: www.cebam.be
- NICE: http://www.nice.org.uk/
- SIGN: http://www.sign.ac.uk/
- CRD: http://www.york.ac.uk/inst/crd/index.htm
- INAHTA: http://www.inahta.org/inahta_web/index.asp
- EUNETHA: http://www.eunethta.net/
- CONSORT statement: http://www.consort-statement.org/
- QUOROM statement: http://www.consort-statement.org/QUOROM.pdf
- GRADE working group: http://www.gradeworkinggroup.org/
- AGREE: http://www.agreecollaboration.org/
- PubMed: http://www.pubmed.gov
- Embase: http://www.embase.com/
- FDA: http://www.fda.gov/
- EMEA: http://www.emea.europa.eu/

KCE Process notes

Searching for Evidence and Critical Appraisal

- Good Clinical Practice (GCP)
- Health Technology Assessment (HTA)
- Health Services Research (HSR)