

## **Searching for Studies**

**Prepared for:**

Agency for Healthcare Research and Quality  
U.S. Department of Health and Human Services  
540 Gaither Road  
Rockville, MD 20850  
[www.ahrq.gov](http://www.ahrq.gov)

**AHRQ Publication No. xx-EHCxxx-EF  
<Month Year>**

Effectiveness and Comparative Effectiveness Reviews, systematic reviews of existing research on the effectiveness, comparative effectiveness, and comparative harms of different medical tests, are intended to provide relevant evidence to inform real-world health care decisions for patients, providers, and policymakers. In an effort to improve the transparency, consistency, and scientific rigor of the work of the Effective Health Care (EHC) Program, through a collaborative effort within the Evidence-based Practice Center Program, have developed a Methods Guide for Medical Test Reviews. We intend that these documents will serve as a resource for our EPCs as well as for other investigators interested in conducting systematic reviews on medical tests.

This Medical Test Methods guide is intended to be a practical guide for those who prepare and use systematic reviews on medical tests. This document complements the EPC Methods Guide on Comparative Effectiveness Reviews (<http://www.effectivehealthcare.ahrq.gov/index.cfm/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productid=318>), which focuses on methods to assess the effectiveness of treatments and interventions. The guidance here applies the same principles for assessing treatments to the issues and challenges in assessing medical tests and highlights particular areas where the inherently different qualities of medical tests necessitate a different or variation of the approach to systematic review compared to a review on treatments. We provide guidance in stepwise fashion for those conducting a systematic review.

The *Medical Test Methods Guide* is a living document, and will be updated as further empirical evidence develops and our understanding of better methods improves. Comments and suggestions on the *Medical Test Methods Guide* and the Effective Health Care Program can be made at [www.effectivehealthcare.ahrq.gov](http://www.effectivehealthcare.ahrq.gov).

None of the investigators have any affiliations or financial involvement that conflicts with the material presented in this report.
---

## Paper 4. Searching for Studies

Locating all published studies relevant to the key questions is a goal of all systematic reviews. Inevitably, Evidence-based Practice Centers (EPCs) encounter variation in whether or how a study is published and in how the elements of a study are reported in the literature or indexed by organizations such as the National Library of Medicine. A systematic search must attempt to overcome these problems to identify all relevant studies, taking into account the usual constraints on time and resources.

With studies of medical tests, locating all available studies is especially important because the results of studies of medical tests themselves tend to be highly variable.<sup>1-2</sup> In the face of such challenges, searches need to use multiple approaches to have high recall (sensitivity). Unfortunately, sensitivity usually comes at the cost of relevance (specificity). Any systematic review of medical tests is likely to involve a good deal of human labor identifying relevant articles from large batches of potentially relevant articles to be sure that none is missed. In this paper, we discuss some of the challenges in identifying studies, focusing on medical tests.

### Common Challenges

Systematic reviews of test strategies for a given condition require a search on each of the relevant test strategies under consideration. In conducting the search, an EPC may use one of two approaches. The EPC may search on all possible tests used to evaluate the given disease, which requires knowing all the possible test strategies available, or the EPC may search on the disease or condition and then focus on medical test evaluation for that disease.

When a review focuses on specific named tests, searching is relatively straightforward. The names of the tests can be used to locate studies, and a specific search for the diagnostic concept may not be necessary.<sup>3-4</sup>

However, searches for a disease or condition are broader searches and greatly increase the burden of work in filtering down to the relevant studies on medical test evaluation.

### Principles for Addressing the Challenges

#### Principle 1: Do not rely on search filters alone

Several search filters (sometimes called “hedgies”), which are pre-prepared and tested searches that can be combined with searches on a particular disease or condition, have been developed to aid systematic reviewers evaluating medical tests. Most of these filters have been developed for MEDLINE®.<sup>1,3-6</sup> In particular, one filter<sup>7</sup> is used in the PubMed® Clinical Queries for diagnosis (Table 4-1). Search filters have also been developed specifically for diagnostic imaging<sup>8</sup> and for EMBASE®.<sup>9-10</sup>

**Table 4-1. Diagnosis Clinical Query for PubMed**

Category	Optimization	Sensitivity/Specificity	PubMed search string
Diagnosis	Sensitivity/breadth	98%/74%	(sensitivity*[Title/Abstract] OR sensitivity and specificity[MeSH Terms] OR diagnos*[Title/Abstract] OR diagnosis[MeSH:noexp] OR diagnostic*[MeSH:noexp] OR diagnosis,differential[MeSH:noexp] OR diagnosis[Subheading:noexp])
	Specificity/narrowness	64%/98%	(specificity[Title/Abstract])

Unfortunately, although these search filters are useful for the casual searcher who simply needs some good articles on diagnosis, they are inappropriate for use in systematic reviews of clinical effectiveness. Several researchers<sup>2,6,11-12</sup> have reported that using these filters for systematic reviews may result in relevant studies being missed. Vincent found that most of the available filters perform better when they are being evaluated than when they are used in the context of an actual systematic review;<sup>12</sup> this finding is particularly true for studies published before 1990 because of non-standardized reporting and indexing of medical test studies.

In recent years, improved reporting and indexing of randomized controlled trials (RCTs) have made such trials much easier to find. There is reason to believe that reporting and indexing of medical test studies will similarly improve in the future.<sup>11</sup> In fact, Kastner and colleagues<sup>13</sup> recently reviewed 22 systematic reviews of diagnostic accuracy published in 2006 to determine whether the PubMed Clinical Queries for diagnosis would be sufficient to locate all the primary studies that the 22 systematic reviews had identified through traditional search strategies. Using these filters in MEDLINE and EMBASE, the authors found 99 percent of the articles in the systematic reviews they examined, and they determined that the missed articles would not have altered the conclusions of the systematic reviews. The authors therefore concluded that filters may be appropriate when searching for systematic reviews of medical test accuracy. However, until more evidence of their effectiveness is found, we recommend that EPCs not rely on them exclusively.

## **Principle 2: Do not rely on controlled vocabulary alone**

It is important to use all known variants of the test name when searching, and these may not all be controlled vocabulary terms. Because reporting and indexing of studies of medical tests is so variable, one cannot rely on controlled vocabulary terms alone.<sup>4</sup>

Because indexing of medical tests is variable, using textwords for particular medical tests will help to identify medical test articles that have not yet been indexed or that have not been indexed properly.<sup>3</sup> Filters may suggest the sort of textwords that may be appropriate. As always—but in particular with searches for studies of medical tests—we advise EPCs to search more than one database and to tailor search strategies to each individual database.<sup>14</sup>

Until reporting and indexing are improved and standardized, a combination of highly sensitive searches and brute force article screening will remain the best approach for systematically searching the medical test literature.<sup>2,6,11-12</sup> Even an initial sensitive search is likely to miss

relevant articles, and following cited references from relevant articles (hand searching) and identifying articles that cite key studies remain important sources of citations.<sup>15</sup>

Because the FDA regulates many medical tests as medical devices, another potential source of information is regulatory documents. Reviewers who know the name of specific tests can search for regulatory documents at the FDA's Device website:

<http://www.accessdata.fda.gov/scripts/cdrh/devicesatfda/index.cfm>.

## Illustration

In the AHRQ report, *Testing for BNP and NT-proBNP in the Diagnosis and Prognosis of Heart Failure*,<sup>16</sup> the medical tests in question were known. Therefore, the search consisted of all possible variations on the names of these tests and did not need to include a search string to capture the diagnostic testing concept. By contrast, in the AHRQ report, *Effectiveness of Noninvasive Diagnostic Tests for Breast Abnormalities*,<sup>17</sup> all possible diagnostic tests were not known. For this reason, the search strategy included a search string meant to capture the diagnostic testing concept, and this relied heavily on textwords. The actual search strategy used in PubMed to capture the concept of diagnostic tests was as follows: diagnosis OR diagnose OR diagnostic OR di[sh] OR "gold standard" OR "ROC" OR receiver operating characteristic" OR sensitivity and specificity[mh] OR likelihood OR "false positive" OR "false negative" OR "true positive" OR "true negative" OR "predictive value" OR accuracy OR precision.

## Summary

Key points are:

- Currently, diagnostic search filters—or, more specifically, the reporting and indexing of medical test studies upon which these filters rely—are not sufficiently well-developed to be depended upon exclusively for systematic reviews.
- If the full range of tests is known, EPCs may not need to search for the concept of diagnostic testing; searching for the specific test using all possible variant names may be sufficient.
- Combining highly sensitive searches utilizing textwords with hand searching and acquisition and review of cited references in relevant papers is currently the best way to identify all or most relevant studies for a systematic review.
- Do not rely on controlled vocabulary alone.
- Check Devices@FDA.

## References

1. Bachmann LM, Coray R, Estermann P, et al. Identifying diagnostic studies in MEDLINE: reducing the number needed to read. *J Am Med Inform Assoc* 2002;9(6):653-8.

2. Leeflang MM, Scholten RJ, Rutjes AW, et al. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol* 2006;59(3):234-40.
3. Deville WL, Bezemer PD, Bouter LM. Publications on diagnostic test evaluation in family medicine journals: an optimal search strategy. *J Clin Epidemiol* 2000;53(1):65-9.
4. van der Weijden T, CJ IJ, Dinant GJ, et al. Identifying relevant diagnostic studies in MEDLINE. The diagnostic value of the erythrocyte sedimentation rate (ESR) and dipstick as an example. *Fam Pract* 1997;14(3):204-8.
5. Haynes RB, Wilczynski N, McKibbin KA, et al. Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *J Am Med Inform Assoc* 1994;1(6):447-58.
6. Ritchie G, Glanville J, Lefebvre C. Do published search filters to identify diagnostic test accuracy studies perform adequately? *Health Info Libr J* 2007;24(3):188-92.
7. Haynes RB, Wilczynski NL. Optimal search strategies for retrieving scientifically strong studies of diagnosis from Medline: analytical survey. *BMJ* 2004;328(7447):1040.
8. Astin MP, Brazzelli MG, Fraser CM, et al. Developing a sensitive search strategy in MEDLINE to retrieve studies on assessment of the diagnostic performance of imaging techniques. *Radiology* 2008;247(2):365-73.
9. Bachmann LM, Estermann P, Kronenberg C, et al. Identifying diagnostic accuracy studies in EMBASE. *J Med Lib Assoc* 2003;91(3):341-6.
10. Wilczynski NL, Haynes RB. EMBASE search strategies for identifying methodologically sound diagnostic studies for use by clinicians and researchers. *BMC Med* 2005;3:7.
11. Doust JA, Pietrzak E, Sanders S, et al. Identifying studies for systematic reviews of diagnostic tests was difficult due to the poor sensitivity and precision of methodologic filters and the lack of information in the abstract. *J Clin Epidemiol* 2005;58(5):444-9.
12. Vincent S, Greenley S, Beaven O. Clinical evidence diagnosis: developing a sensitive search strategy to retrieve diagnostic studies on deep vein thrombosis: a pragmatic approach. *Health Info Libr J* 2003;20(3):150-9.
13. Kastner M, Wilczynski NL, McKibbin AK, et al. Diagnostic test systematic reviews: bibliographic search filters ("Clinical Queries") for diagnostic accuracy studies perform well. *J Clin Epidemiol* 2009;62(9):974-81.
14. Honest H, Bachmann LM, Khan K. Electronic searching of the literature for systematic reviews of screening and diagnostic tests for preterm birth. *Eur J Obstet Gynecol Reprod Biol* 2003;107(1):19-23.

15. Whiting P, Westwood M, Burke M, et al. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *J Clin Epidemiol* 2008;61(4):357-364.
16. Balion C, Santaguida P, Hill S, et al. Testing for BNP and NT-proBNP in the Diagnosis and Prognosis of Heart Failure. Evidence Report/Technology Assessment No. 142. (Prepared by the McMaster University Evidence-based Practice Center under Contract No. 290-02-0020). AHRQ Publication No. 06-E014. Rockville, MD: Agency for Healthcare Research and Quality. September 2006. Available at: [www.ahrq.gov/downloads/pub/evidence/pdf/bnp/bnp.pdf](http://www.ahrq.gov/downloads/pub/evidence/pdf/bnp/bnp.pdf). Accessed April 16, 2009.
17. Bruening W, Launderers J, Pinkney N, et al. Effectiveness of Noninvasive Diagnostic Tests for Breast Abnormalities. Comparative Effectiveness Review No. 2. (Prepared by ECRI Evidence-based Practice Center under Contract No. 290-02-0019.) Rockville, MD: Agency for Healthcare Research and Quality. February 2006. Available at: <http://effectivehealthcare.ahrq.gov/repFiles/BrCADx%20Final%20Report.pdf>. Accessed April 16, 2009.

DRAFT