---

**Article:**
Berry, E., Kelly, S., Hutton, J. et al. (2000) Identifying studies for systematic reviews - An example from medical imaging. International Journal of Technology Assessment in Health Care, 16 (2). pp. 668-672. ISSN: 1471-6348

https://doi.org/10.1017/S0266462300101229

---

**International Journal of Technology Assessment in Health Care**, **16:2** (2000), 668–672.


# IDENTIFYING STUDIES FOR SYSTEMATIC REVIEWS

## *An Example from Medical Imaging*

**Elizabeth Berry**
**Steven Kelly**
*University of Leeds*

**John Hutton**
*MEDTAP International Inc.*

**Keith M. Harris**
*Leeds Teaching Hospitals NHS Trust*

**Michael A. Smith**
*University of Leeds*

Abstract

**Objectives:** To determine if published figures on the proportion of articles included in systematic reviews and identified in electronic databases are applicable to an example from medical imaging.
**Methods:** A systematic review was performed. Additionally, sensitivity and precision of a MEDLINE search were compared with values from three published searches, each customized for a specific field.
**Results:** All articles included in the systematic review were in electronic databases. The MEDLINE search had low precision compared with searches in other fields.
**Conclusions:** In a specific area of medical imaging, electronic databases, including MEDLINE, are reliable sources of articles.

**Keywords:** MEDLINE, Online systems, Evaluation studies, Diagnosis

The finding that electronic databases such as MEDLINE may return only about half of the papers found from hand-searching for randomized controlled trials (RCTs) (3;4) is widely cited, and further studies have demonstrated the difficulties of searching for specific study types within (1;9) or across (10) subject areas. Similarly, there can be problems in identifying

This work was carried out with the financial support of the Secretary of State for Health under the U.K. National Health Service Health Technology Assessment Programme, project 94/44/3. The views and opinions expressed do not necessarily reflect those of the Secretary of State for Health. In part, this work was undertaken by the Leeds Teaching Hospitals NHS Trust, which received funding from the National Health Service (NHS) Executive; the views expressed in this publication are those of the authors and not necessarily those of the NHS Executive. The participation of P. Roderick, J. Cullingworth, L. Gathercole, P. J. O'Connor, and J. C. Boyce as members of the multidisciplinary review team is gratefully acknowledged.

relevant studies in fields where there is no strong specialist focus. For example, recently Hobbs et al. (6) performed a systematic review of "near patient testing." Near patient tests are those investigations performed in a clinical setting without the need to send a sample to a remote laboratory. Clearly such tests may arise in a number of clinical specialties, the published literature will be diffuse, and selection of keywords for literature searches difficult. Indeed, it was reported (8) that 24% of the references included in the systematic review of near patient testing, not necessarily RCTs, would have been missed without the input of external experts. However, the usefulness of MEDLINE varies from field to field (7), depends on the type of study sought, and may change over time as indexing changes. Our aim in this work is to demonstrate the different outcomes that can result from the varied searching approaches that are necessary for different subject areas. We have compared the results of a search performed as part of a systematic review in medical imaging (5), which covered a range of study designs and was in a well-defined topic area, with the previously published searches (4;6). We do not perform comparisons by applying previously published search strategies to different topic areas. It would be inappropriate to apply, for example, an RCT search (4) to near patient testing (6), as in the latter case a range of different study designs is applicable. Instead, we present results about the ability of the reported MEDLINE strategy to retrieve articles compared with the gold standard search in that field, to show that conclusions about the usefulness of MEDLINE should not be applied indiscriminately across subject areas.

## METHODS

A systematic review was undertaken for the U.K. National Health Service (NHS) Health Technology Assessment (HTA) R&D Programme in 1996–97, concerning the use of endoscopic ultrasound (EUS) in gastroesophageal cancer. Harris et al. (5) have published a comprehensive report of this review elsewhere. To address a range of questions about the use of EUS, from its effect upon patient outcome in any clinical area to its staging performance for primary and lymph node tumors of the esophagus, stomach, or at the cardia, a very broad search strategy was used. Electronic searches were conducted of MEDLINE, Institute of Scientific Information (ISI), Cochrane Library, EMBASE, Inside Information Plus (2), FirstSearch (11), and System for Information on Grey Literature in Europe (SIGLE). Specially designed search strategies (5) were used to retrieve information from MEDLINE and ISI from 1981 to 1996 inclusive. The first search strategy used a comprehensive combination of technological free-text words and was refined as new words were identified. To identify studies on therapeutic impact, patient outcome, or health economics, the electronic search was refined to include words relating to the type of result reported, for example, those expected in a cost-effectiveness study. The search was not designed to identify study types such as the RCT. The remaining databases were searched using their simpler search interfaces and the words "endoscopic ultrasound," "endoscopic ultrasonography," "EUS," and "endosonography." Study type was not specified. Citation lists of retrieved papers were hand-searched, and equipment manufacturers, authors, and an electronic mail discussion list contacted for references and unpublished data. A dedicated gastrointestinal radiologist on the multidisciplinary review team, in consultation with gastroenterologists and gastrointestinal surgeons, gave advice at an early stage to focus the search and identify relevant journals and conferences.

Appraisal of the results of the searches was performed: articles in the English language, reporting original data for more than 10 patients, relevant to the questions addressed by the review, and published before January 1997 were included in the review of EUS (5). For studies of staging performance and impact, additional inclusion criteria were applied to ensure that studies were designed to investigate performance or impact, to limit the

anatomical application to the stomach, esophagus, or cardia, and the type of disease to squamous cell carcinoma or adenocarcinoma. Studies of staging performance had to have the gold standard of pathology or histology, present sufficient raw data for completion of a contingency table, and include at least basic information about study design. The results of the staging performance search and appraisal are used in the comparison of searches in this paper.

To highlight the differences in our approach from those of other authors, we compared the performance of our (5) MEDLINE search strategy for studies of staging performance (search 4) with two searches described by Dickersin et al. (4) (search 1 and search 2), and one by Hobbs et al. (search 3) (6;8). The gold standard set of papers in each case included papers that were identified by the authors using their complete search methodology and that subsequently satisfied appraisal for inclusion in their review. We have assumed that all the authors applied a sufficiently rigorous methodology to identify all relevant papers, justifying the use of this set as a gold standard. Four performance parameters were defined:

1. MEDLINE search sensitivity = (number of papers in gold standard set returned by MEDLINE search/number of papers in gold standard set) × 100%;
2. MEDLINE search precision = (number of papers in gold standard set returned by MEDLINE search/number of papers retrieved from MEDLINE before appraisal) × 100%;
3. MEDLINE-indexed search sensitivity = (number of papers in gold standard set returned by MEDLINE search/number of papers in the gold standard set known to be indexed in MEDLINE) × 100%; and
4. Inclusion rate of MEDLINE-indexed papers = (number of papers in gold standard set indexed in MEDLINE/number of papers in gold standard set) × 100%.

## RESULTS

Over the whole EUS review, 94% (44 from 47) of the studies included were from the MEDLINE database, and the corresponding figure was 92% (33 from 36) for the staging performance review. Two of the three papers not in MEDLINE were identified electronically in one or more of the other databases searched. One paper was identified only from hand-searching of citation lists; it was not listed in any of the electronic databases.

Results of the comparison of published MEDLINE searches appear in Table 1.

## DISCUSSION

In Table 1 it can be seen that the other reviews included a larger percentage of papers indexed in MEDLINE than did our EUS review. The difficulty in defining an adequate search strategy resulted in a MEDLINE-indexed search sensitivity of under 50% in searches 1 and

**Table 1.** Comparison of Published MEDLINE Search Strategies

| | MEDLINE search sensitivity (%) | MEDLINE search precision (%) | MEDLINE-indexed search sensitivity (%) | Inclusion rate of MEDLINE-indexed papers (%) |
|---|---|---|---|---|
| Search 1 (4) | 44.5 | 47.9 | 47.3 | 94.1 |
| Search 2 (4) | 81.8 | 12.7 | 86.9 | 94.1 |
| Search 3 (6) | 19.6 | 6.7 | 20.2[a] | 95.1[a] |
| Search 4 (5) (EUS search) | 91.7 | 1.2 | 100 | 91.7 |

[a] MEDLINE availability was determined by us in January 1999 and is not taken from the published review.

3, while our search gave 100% retrieval from MEDLINE of papers known to be indexed in MEDLINE. It is obvious, however, that this has been achieved at the expense of precision, and our search differs markedly from those of earlier published reports having a very low precision and high sensitivity.

The principal difficulty with MEDLINE is not the omission of journals from MEDLINE indexing, but the successful retrieval of relevant studies. The task is more difficult when the search is for certain study designs (4) or across a range of fields (6;8) than for a specific subject area. In our review a highly sensitive search strategy was used, and the strategy was improved iteratively by adding newly identified free-text and MeSH terms to increase sensitivity. This type of search strategy has been advocated by Jadad and McQuay (7) and by van der Wiedjen and colleagues (12), who obtained high sensitivity even in the diverse field of primary care. This low precision methodology, however, is very labor intensive and for some reviews may not be justifiable.

Although a slightly smaller percentage of our gold standard set of studies was indexed in MEDLINE than for the other authors, we were able to find all but one electronically. This may be because the topic was associated with a high-technology device, and information was published in journals rather than in books and other less available literature. None of our included papers was from before 1988, and so no historical searching prior to indexing in the electronic databases was necessary. Although widespread canvassing of expert opinion was avoided at the search stage, expert input was useful to help identify work in progress and essential for the critical appraisal of retrieved studies. In the EUS field, although we were able to devise a successful MEDLINE search strategy, searching of MEDLINE alone still returned an incomplete list of papers. Eight percent of the studies fulfilling the inclusion criteria for our review were not in MEDLINE but were found in other databases and by hand searches.

One of the basic premises of the systematic review is that it involves an exhaustive search. In a subject such as this area of medical imaging where study results are published primarily in journals, we have shown that an extensive search of electronic databases is an effective methodology. It is emphasized that equally valid search strategies can give widely differing retrieval rates, depending on the nature of the literature sought. Just as results for the previous studies were found to differ from those of our EUS review, neither are our EUS results universally applicable.

## REFERENCES

1. Bender, J. S., Halpern, S. H., Thangaroopan, M., Jadad, A. R., & Ohlsson, A. Quality and retrieval of obstetrical anaesthesia randomized controlled trials. *Canadian Journal of Anaesthesiology*, 1997, 44, 14–18.
2. British Library: Inside information plus. Available at: http://www.bl.uk/online/inside. 1998.
3. Chalmers, I., Dickersin, K., & Chalmers T. C. Getting to grips with Archie Cochrane's agenda. *British Medical Journal*, 1992, 305, 786–88.
4. Dickersin, K., Scherer, R., & Lefebvre, C. Systematic reviews: Identifying relevant studies for systematic reviews. *British Medical Journal*, 1994, 309, 1286–91.
5. Harris, K. M., Kelly, S., Berry, E., et al. Systematic review of endoscopic ultrasound in gastro-oesophageal cancer. *Health Technology Assessment*, 1998, 2(18).
6. Hobbs, F. D. R., Delaney, B. C., Fitzmaurice, D. A., et al. A review of near patient testing in primary care. *Health Technology Assessment*, 1997, 1(5).
7. Jadad, A. R., & McQuay, H. J. Be systematic in your searching. *British Medical Journal*, 1993, 307, 66.
8. McManus, R. J., Wilson, S., Delaney, B. C., et al. Review of the usefulness of contacting other experts when conducting a literature search for systematic reviews. *British Medical Journal*, 1998, 317, 1562–63.

9. Marson, A. G., & Chadwick, D. W. How easy are randomized controlled trials in epilepsy to find on Medline? The sensitivity and precision of two Medline searches. *Epilepsia*, 1996, 37, 377–80.
10. Neal, B., Rodgers, A., Mackie, M. J., & MacMahon, S. Forty years of randomised trials in the New Zealand Medical Journal. *New Zealand Medical Journal*, 1996, 109, 372–73.
11. Online Computer Library Center. Available at: http://www.oclc.org. 1998.
12. van der Weijden, T., Ijzermans, C. J., Dinant, G. J., et al. Identifying relevant diagnostic studies in Medline: The diagnostic value of the erythrocyte sedimentation rate (ESR) and dipstick as an example. *Family Practice*, 1997, 14, 204–08.