Developing High-Quality Search Strategies for Systematic Reviews

A Guide to the Galaxy of Search Terms, Electronic Databases, and More

Shari Bolen MD, MPH[1]
Nisa Maruthur MD, MHS[2]
Sonal Singh MD, MPH[2]
Lisa Wilson MPH[2]
Eric Bass MD, MPH[2]

[1] MetroHealth/Case Western Reserve University, Cleveland, Ohio
[2] Johns Hopkins University, Baltimore, Maryland

Background

In the world of EBM & comparative effectiveness, systematic reviews are increasingly popular

- They can help determine clinical practice
- They can help identify research gaps

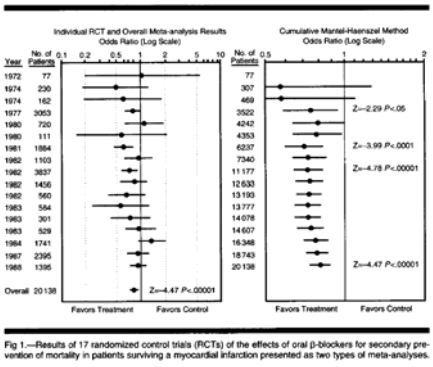

Fig 1.—Results of 17 randomized control trials (RCTs) of the effects of oral β-blockers for secondary prevention of mortality in patients surviving a myocardial infarction presented as two types of meta-analyses.

Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. *JAMA*. 1992;268(2):240-248.

Background

- BUT not all are free of bias
- The search strategy is crucial to inclusion of articles which will influence your conclusions
- Today we plan to discuss steps YOU can take to produce a high-quality bias-free search strategy

Objectives

We will use an example case to:

- Describe the types of search terms & databases useful when conducting a systematic electronic search for relevant studies
- Design a search strategy using PubMed
- Translate a PubMed strategy to other electronic databases, such as EMBASE and Cochrane
- Explain the role of manual searching
- Describe the pros and cons of searching for non peer-reviewed literature

Example case

- Compare the effectiveness and safety of oral medications to treat type 2 diabetes
- **PICOT S** (already defined for you):
  - **P**=population: type 2 diabetic adults
  - **I**=interventions: oral diabetes medications (monotherapy and dual combination therapy)
  - **C**=comparisons: head to head and indirect
  - **O**=outcomes: Hgba1c, cvd morbidity and mortality, microvascular complications, adverse events
  - **T**=timing: studies have to be > 3 months for all outcomes except safety (> 1 month)
  - **S**=setting or study design: outpatient primary care or specialty clinics for setting; Study design - RCTs for intermediate outcomes and observational studies with comparison groups for other outcomes

Developing the search strategy in PubMed

- Enter the search term in PubMed
- Go to Details to see terms actually searched
- Look for associated MESH term
- Look at year MESH introduced
- See if any prior MESH terms need to be added to the search strategy
- Search the term as keyword under title and abstract [tiab] and search appropriate MESH terms

Developing the search strategy in PubMed

- Add other synonymous terms
- Truncate words if appropriate by using *
  - Hypoglycem* will give hypoglycemia and hypoglycemic terms
- Example: glyburide[tiab] OR glyburide[mh] OR glibenclamide[tiab] OR "hypoglycemic agents"[mh] OR "sulfonylurea compounds"[mh] OR sulfonylurea*[tiab] OR sulphonylurea*[tiab]

Developing the search strategy in PubMed

- If a term has multiple words like type 2 diabetes, may want to do several iterations
- For instance: ("type 2" and diabetes) or "type 2 diabetes" or etc…
- PubMed will then search for both type 2 and diabetes terms without requiring the terms to come right after one another
- Combine concept terms with AND. Type 2 diabetes terms AND medications terms.

Limiting the search strategy in PubMed

When appropriate, can limit entire search to:

- randomized trials
- humans
- publication date
- english language
- original data (excluding reviews and editorials)

Be careful/avoid using **NOT in the search strategy ("*NOT animals*" would exclude studies that had both humans and animals)

Checking your search strategy

- Use one half of key articles to help identify search terms
- Use the other half of your key articles to check your search strategy
- If missing key articles, determine what additional terms may need to be added
- Related articles search – good to help find other key articles ONLY

Choosing the best search strategy

- Go to history to see the number of citations for several searches you have developed
- Choose the search strategy:
  - with all your key articles
  - as sensitive as possible (with a reasonable work load)

Translating to EMBASE

- 7,000 journals (1,800 more biomedical titles compared to MEDLINE)
  - More non-English language journals
  - Emphasis on pharmacologic research
- Emtree – 2x as many terms (vs MeSH)
- Example: glyburide:ti,ab OR glyburide/exp OR glibenclamide:ti,ab OR 'hypoglycemic agents'/exp OR 'sulfonylurea compounds'/exp OR sulfonylurea*:ti,ab OR sulphonylurea*:ti,ab

Translating to Cochrane

Consists of 6 databases including:

- Cochrane Database of Systematic Reviews-Useful for finding SRs/MA on this topic and reviewing included/excluded trials
- Cochrane Central Register of Clinical trials for RCTs
- Additional databases such as Database of Reviews of Effectiveness may not be that helpful
- Unlike EMBASE can using same MESH terms in Cochrane
- MESH descriptor Glyburide exploded in Cochrane Central Register of Controlled Trials yields 433 of 608405 records
- Can use NEAR or ADJ to search two terms close together such as "type 2" NEAR "diabetes"

Grey Literature-Challenges

- Difficult to choose whether to search FDA documents, trials registries, PHRMA or manufacturers databases
- Creates challenges for Study Flow sheet as studies identified differently than peer reviewed articles so may get duplication.
- Not enough methodological detail to adequately assess bias
- Outcomes may be reported in different formats (analytical challenges)
- Resource intensive and yield uncertain for efficacy outcomes.
  - However may potentially yield data on adverse effects, especially FDA reviews

Manual searching

- Hand searching is one way to make sure you do not miss any articles
- Includes scanning journals that may not be in PubMed or not up to date within PubMed
- Includes checking references of included articles
- Includes checking references of review articles
- Query experts for missing relevant articles

Summary Points

- Query experts on key articles, and ask them to help refine your question using PICOTS
- Choose appropriate electronic databases
- Develop a search strategy in PubMed using your clinical knowledge and key articles
- Avoid using outcome terms in the search strategy
- Avoid adding limitations to the search strategy

Summary Points

- Translate the strategy to other databases
- Conduct a thoughtful manual search
- Document dates you searched each database
- Have experts review final articles to ensure no articles are missing

Small group exercises

Example case:

A. Compare the effectiveness and safety of oral medications to treat type 2 diabetes

B. **PICOT S** (already defined for you):
P=population: type 2 diabetic adults
I=interventions: FDA-approved oral diabetes medications (monotherapy and dual combination therapy)
C=comparisons: head to head and indirect
O=outcomes: Hgba1c, cvd morbidity and mortality, microvascular outcomes , adverse events
T=timing: studies have to be > 3 months for all outcomes except safety (> 1 month)
S=setting or study design: setting was outpatient primary care or specialty clinics; study design was randomized and controlled clinical trials along with observational studies with a comparison group for adverse events and long term outcomes, and randomized controlled trials for intermediate outcomes such as HgbA1c.

Participatory section questions:

Question 1a. What databases might you search?

Question 1b. What about searching for "grey" or non-peer-reviewed literature?

(Spend 5-10 minutes on these questions)

Answer to 1a and 1b:

For this specific case, PubMed, EMBASE, and Cochrane Library are probably sufficient since they will likely capture most of the articles related to diabetes medications. Additional databases would increase the sensitivity of your search, but decrease specificity and likely not add anything that would markedly change the results for this particular question.

If you had a topic on education, you might consider searching ERIC (educational database). If you had a topic dealing with adherence or behavioral interventions, you might consider searching CINAHL (nursing and allied health), since they would have a lot of literature on adherence or behavioral interventions. If you had a cancer topic, searching CancerLit would make sense. See below for additional databases to consider depending on your topic. Consulting with a librarian at your institution may help you determine which databases to search for your particular question.

"Grey" literature searches (see list of sources below) make some sense for this specific case, especially since pharmaceutical companies may work hard to avoid showing less benefit from their

drug versus another drug for the same disease and may try to avoid showing harm. However, one must be careful to only include the grey literature with little risk of bias. One would want to search the FDA website and the government supported clinical trials registry website for instance, but not necessarily rely on data found on a pharmaceutical website since the company may only be listing a few of their studies there which might bias your results to be more favorable to their drug. Also, it may be difficult to assess quality of these studies since many times methods are not well reported. It is also difficult to tell whether there are duplicates of peer-reviewed literature in these databases since authors are not mentioned making analysis of results challenging. We would not recommend searching for dissertations or meeting abstracts since these results many times may be different than what gets published in the peer-reviewed literature although this could be debated.

Electronic databases with peer-reviewed articles:

1. PubMed or Medline – free health literature database through National Library of Medicine. Would always search this database in a systematic review.
2. EMBASE – your institution must pay to have this database available for searching and is expensive. Has more European journals indexed than PubMed, and has more pharmaceutical literature. In a study comparing Medline with EMBASE on three common topics (low back pain, rheumatoid arthritis, and osteoporosis), one-third of the controlled clinical trials were identified by only one of these databases (see reference list). Hence the importance of including both when possible.
3. Cochrane Library – contains high-quality standardized systematic reviews done by the Cochrane Collaboration. Also has a searchable database for controlled clinical trials. Would include in almost all systematic reviews.
4. CINAHL – This is a nursing and allied health database.
5. ERIC – Educational database.
6. CancerLit – Cancer literature database.
7. PsychInfo – Psychiatric literature database.
8. There are probably several others. It is helpful to consult with people from other fields if the question is not well-linked to one of the above databases.

"Grey" or non-peer-reviewed literature sites:

1. Google – only has access to items that are publicly available
2. FDA website – important for drug studies

3. Pharmaceutical websites – can be dangerous to use these as the pharmaceutical company may only make public beneficial studies which could bias your results.

4. Meeting abstracts – may be problematic since many people discuss preliminary results at meetings which may change once a full analysis is completed.

5. Clinical Trials Registry - This is a publicly available website where investigators must report results from their clinical trial and would be an important source for a systematic review.

6. Dissertations – Again, things may change once peer-reviewed so would likely avoid searching for this.

Question/activity 2: Brainstorm possible search terms for the example case and then share with the group (Spend 10 min discussing)

***Answer to 2*:**

1. Intervention or medication terms: januvia, thiazolidinediones, DPP-IV inhibitors, rosiglitazone, pioglitazone, sitagliptan, metformin, glucophage, glyburide, glipizide, oral hypoglycemic drugs, acarbose, miglitol, nateglinide, repaglinide, sulfonylureas, glibenclamide, glucovance, etc...

- Would include glibenclamide since it has the same chemical formulation as glyburide.

- Would include generic and brand names of drugs.

2. Population of interest terms: type 2 diabetes or type II diabetes or NIDDM

3. Would avoid searching for terms related to the outcome of interest since many times we are evaluating outcomes that may appear in the article but not the abstract especially if they showed no effect on that outcome.

4. It is important to include as many synonyms for search terms as possible to improve sensitivity, but you need to balance this with the need for some specificity in order to have a reasonable workload. This will require some testing of different search strategies using your key articles.

5. Key articles are usually critical articles that you know would be included in a systematic review of your topic. You can obtain these through searching PubMed and querying experts or searching references of review articles on your topic. Typically, we get 8-10 key articles and use half of them to determine good search terms that we might be missing, and use the other half to make sure our search strategy is capturing these important articles. If the search strategy is missing some of the articles, then you would need to determine why – usually a missing search term.

Question 3. How can you limit the numbers of titles using the search strategy without introducing bias? (Spend 5-10 min discussing)

***Answer Key to question 3*:**

a. Can you limit by study design in your search strategy itself? There are validated search strategies that help you limit to randomized controlled trials or controlled clinical trials (see reference list) assuming this makes sense given your question and amount of time/people you have to conduct a systematic review. In the case example, we did not limit the search strategy itself by study design since we wanted observational studies for adverse effects and long term outcomes. We excluded observational studies during abstract and article review for the intermediate outcomes such as HgbA1c since we had enough randomized trial data to answer that question.

b. Can you limit by publication date? Only makes sense if the drugs of interest or intervention were introduced after a specific date.

c. Can you limit to humans? Yes, but you must be careful with the wording. Using **NOT** in PubMed should generally be avoided since writing NOT animals will exclude trials with humans and animals. Need to write "NOT (animal[mh] NOT human [mh])" in PubMed.

d. Can you limit to specific languages? Depends on the question topic and what journal you are targeting for eventual publication. In the case example, we could limit to English language since we were focused on FDA approved medications used in the United States. Cochrane instructs people to NOT limit by English language in the search strategy. They recommend you document any articles excluded based on language as part of your title and abstract review. It will be harder to publish in Lancet or BMJ if limit by language since these are international journals. You may also get "dinged" by reviewers if you limit by English language for an international topic.

e. Can you limit by article type (i.e. reviews, comments, letters, editorials)? You could limit the search strategy to exclude reviews or editorials or comments but could include letters since many letters have original data. Can do a specific systematic review search to identify critical recent systematic reviews in order to search their reference list for critical references as well as to make sure there is a need for a systematic review.

f. I want to emphasize that we can exclude based on the above considerations during title and abstract review and does NOT need to be part of the search strategy. Only beneficial to limit the search strategy if have too large a workload and/or do not jeopardize the sensitivity of the review in a way that might introduce bias or change results.

Search strategies from the published literature – Are they high quality? Why or why not?

Example 1: A systematic review evaluating obesity and its' association with cervical cancer screening.

> They used terms for obesity AND cervical cancer AND screening in their search strategy. Searched PubMed, CINAHL, and Cochrane. They included dates databases were searched and how far back they searched them. They conducted a manual search of references of included articles and reviews, and used key articles to develop the search strategy. Search strategy was limited to humans and English language articles. Appendix has complete list of terms used in the search strategy.

Example 1 Answer: By using only obesity terms, this may have biased the authors to find only positive studies (publication bias). They could have looked at studies evaluating any factor associated with cervical cancer screening although this would have expanded the titles to a possibly unmanageable workload. Also, they may have missed articles since they did not search CancerLit or EMBASE. It is hard to know if that would have introduced a bias or not. EMBASE was not used since the authors were interested in factors associated with cervical cancer screening in the U.S., but could have similar factors associated with screening in other developed countries with similar screening practices.

Example 2: A systematic review evaluating NSAID effects on cognition. Article states that they searched MEDLINE using the search term "NSAID" combined with "Alzheimer's disease," "dementia," or "cognitive decline," limited to studies in humans. They did not apply any language restriction.

Example 2 answer: By using only NSAIDs, they may have missed relevant articles of specific NSAIDs. They should have included each individual drug name in their search strategy. They may not have had enough terms for cognition, although difficult to tell since they had no appendix indicating details of their search strategy. While not using a language restriction, they did limit to only searching Medline and not Embase or Cochrane Library which may have caused them to miss some articles. Lastly, there is no record of dates the databases were searched. This could have biased their outcome in a number of ways.

References

Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. BMJ 1994; 309: 1286-1291

Golder S, McIntosh HM, Duffy S, Glanville J, Centre for Reviews and Dissemination and UK Cochrane Centre Search Filters Design Group. Developing efficient search strategies to identify reports of adverse effects in MEDLINE and EMBASE. Health Information and Libraries Journal 2006; 23: 3-12.

Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ 2005; 331: 1064-1065.

Sampson M, McGowan J. Errors in search strategies were identified by type and frequency. Journal of Clinical Epidemiology 2006; 59: 1057-1063.

Wieland S, Dickersin K. Selective exposure reporting and Medline indexing limited the search sensitivity for observational studies of the adverse effects of oral contraceptives. J Clin Epidemiol. 2005 Jun;58(6):560-7.

The Cochrane Handbook. Chapter 6: Searching for studies. Authors: Carol Lefebvre, Eric Manheimer and Julie Glanville on behalf of the Cochrane Information Retrieval Methods Group. Online at http://www.cochrane-handbook.org/. Accessed 03/25/2010.

Suarez-Almazor M, Belseck E, Homik J, Dorgan M, Ramos-Remus C. Identifying clinical trials in the medical literature with electronic databases: MEDLINE alone is not enough. Controlled Clinical Trials 21:476–487 (2000).

Robinson K, Dickersin K. Development of a highly sensitive search strategy for the retrieval of reports of controlled trials using PubMed. International Journal of Epidemiology 2002; 31: 150-153.