Literature search methodology for a systematic review

APHA 2015
LI 1008.0: Systematic Reviews
Chicago, IL

Barbara Folb, MM MLS MPH
folb@pitt.edu
Presenter Disclosures

Barbara L Folb

(1) The following personal, professional, or financial relationships with commercial interests relevant to this presentation existed during the past 12 months:

No Relationships to disclose
Learning Objectives

- Differentiate among the many databases available and assess which are the most appropriate databases for the topic.
- Define grey literature and identify where to locate it.
- Explain where to find and how to use search filters.
- Describe the elements needed for an appropriate search strategy.
Assumptions

- After this session
  - You will understand the search process
  - You will not be an expert searcher
  - You will be motivated to work with a librarian for your next systematic review search
Standard 3.1—Conduct a comprehensive systematic search for evidence

- 3.1.1 Work with a librarian or other information specialist trained in performing systematic reviews to plan the search strategy
- 3.1.3 Use an independent librarian or other information specialist to peer review the search strategy

Peer review standard: PRESS checklist (Sampson et al 2008)
3 Search Standards Compared

From IOM, *Finding What Works in Health Care*… Appendix E, Table E-1, pg 266
Search tasks within the Systematic Review Workflow
Cochrane Review Timeline

**Box 2.3.b: Timeline for a Cochrane review**

<table>
<thead>
<tr>
<th>Month</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 – 2</td>
<td>Preparation of protocol</td>
</tr>
<tr>
<td>3 – 8</td>
<td>Searches for published and unpublished studies.</td>
</tr>
<tr>
<td>2 – 3</td>
<td>Pilot test of eligibility criteria.</td>
</tr>
<tr>
<td>3 – 8</td>
<td>Inclusion assessments.</td>
</tr>
<tr>
<td>3</td>
<td>Pilot test of 'Risk of bias' assessment.</td>
</tr>
<tr>
<td>3 – 10</td>
<td>Validity assessments.</td>
</tr>
<tr>
<td>3</td>
<td>Pilot test of data collection.</td>
</tr>
<tr>
<td>3 – 10</td>
<td>Data collection.</td>
</tr>
<tr>
<td>3 – 10</td>
<td>Data entry.</td>
</tr>
<tr>
<td>5 – 11</td>
<td>Follow up of missing information.</td>
</tr>
<tr>
<td>8 – 10</td>
<td>Analysis.</td>
</tr>
<tr>
<td>1 – 11</td>
<td>Preparation of review report.</td>
</tr>
<tr>
<td>12 –</td>
<td>Keeping the review up-to-date.</td>
</tr>
</tbody>
</table>

From the *Cochrane Handbook for Systematic Reviews of Interventions*
Protocol development searching

Phase I
Who is doing what during protocol development

**Investigators**
- Creating protocol using PRISMA-P guidance
- **Delivering exemplar articles on SR topic to librarian**

**Librarian**
- Searching for existing SRs, SR protocols on topic
- Developing preliminary search for protocol
- **Identifying concept terminology, testing it for inclusion**

**Joint tasks**
- Choosing databases to search
- Choosing grey literature sources to search
- Choosing search filters, if needed
Search Related PRISMA-P Items

- Describe the rationale for the review in the context of what is already known
- Provide an explicit statement of the question(s) the review will address with reference to participants, interventions, comparators, and outcomes (PICO)
- Specify the study characteristics (such as PICO, study design, setting, time frame) and report characteristics (such as years considered, language, publication status) to be used as criteria for eligibility for the review
- Describe all intended information sources (such as electronic databases, contact with study authors, trial registers or other grey literature sources) with planned dates of coverage
- Present draft of search strategy to be used for at least one electronic database, including planned limits, such that it could be repeated
Why Exemplar Articles?

- Librarian will use them for

1. Vocabulary source
   - Cool tools to help extract terms:
     - PubReminer [http://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi](http://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi)
     - Yale MeSH Analyzer [http://mesh.med.yale.edu/](http://mesh.med.yale.edu/)
   - The old fashioned way
     - Print title, abstract, MeSH, highlight candidate synonyms

2. Testing searches
   - Does the search being tested retrieve the exemplar articles that are in the database?
Choosing databases
Which database to start in?

- OVID MEDLINE or PubMed?

- Content essentially the same

- Search interfaces different
  - Mostly allow the same functions, different look and feel
  - Each has a few things it does better than the other

- OVID fee-based subscription, PubMed free
Choosing other databases

- Ask your librarian
- Consult systematic review manuals
- Examine existing systematic reviews on topic
  - Cochrane reviews
  - Others with good structured abstracts
  - Published in your target journal for publication
# Systematic Review Database Standards

<table>
<thead>
<tr>
<th>Manual</th>
<th>Database Choice Advice</th>
</tr>
</thead>
</table>
| Cochrane Collaboration. *Cochrane Handbook for Systematic Reviews of Interventions* | • Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and EMBASE (if available to authors) must be searched.  
• As needed include: national, regional databases; subject specific databases; citation indexes; dissertation databases |
| Centre for Reviews and Dissemination. *CRD’s guidance for undertaking reviews in Health care* | • Most use MEDLINE, Embase, and Cochrane Register of Clinical Trials  
• Others depend on topic  
• **Not possible to set standard for number of databases** |
| Joanna Briggs Institute                                                  | • Start with Medline and CINAHL  
• Choose others based on the topic |
Search Guidance, IOM *Finding What Works*

pg 106-107

- Standard 3.1—Conduct a comprehensive systematic search for evidence
  - 3.1.4 Search *bibliographic databases*
  - 3.1.5 Search *citation indexes*
  - 3.1.6 Search *literature cited by eligible studies*
  - 3.1.8 Search *subject-specific databases* if other databases are unlikely to provide all relevant evidence
  - 3.1.9 Search *regional bibliographic databases* if other databases are unlikely to provide all relevant evidence
Database Choice Criteria

- What’s in it?
- How does it work?
- Systematic review standards
- Do you have access?
Criteria: Database Content

- Topics covered
- Type of literature included
  - Journals, books, book chapters, conference proceedings, etc
- Years included
- Fields included for each record
- Indexing, other additions to the published content
Criteria: Functionality

- **Searching options**
  - Operators allowed
  - Phrase searching
  - Similar references search
  - Display, navigation
  - Ease of building, testing search

- **Results management**
  - Save and reuse searches
  - Save citations
  - Export citations
  - Save search strategy for documentation
SR Example: Many databases

From Abstract: ... IN July 2010 we searched the following electronic databases: EMBASE, MEDLINE, PsycINFO, Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, LILACS, Web of Science, Web of Social Science, NLM Gateway (supplemented by a manual search of the most recent abstracts not included in the Gateway database)...

Interventions to improve adherence to antiretroviral therapy in children with HIV infection.
Bain-Brickley D1, Butler LM, Kennedy GE, Rutherford GW.
SR search example: Minimal number of databases

From Abstract:

We searched online bibliographic databases, including MEDLINE and EMBASE, using systematic criteria.


A systematic review of pediatric adherence to antiretroviral therapy in low- and middle-income countries.

Vreeman RC¹, Wiehe SE, Pearce EC, Nyandiko WM.
Is more always better?

- Do you have the resources to finish the job?
- Are extra databases adding articles that will meet inclusion criteria?
Choosing Grey Literature Sources
IOM Search Guidance

- Required elements:
  - 3.2.1 Search grey-literature databases, clinical trial registries, and other sources of unpublished information about studies
  - 3.2.2 Invite researchers to clarify information related to study eligibility, study characteristics, and risk of bias
  - 3.2.3 Invite all study sponsors to submit unpublished data, including unreported outcomes, for possible inclusion in the systematic review
  - 3.2.4 Handsearch selected journals and conference abstracts
  - 3.2.5 Conduct a web search
  - 3.2.6 Search for studies reported in languages other than English if appropriate
What is grey literature?

- Reports
- Conference Abstracts
- Dissertations & Theses
- Registered Trials
- Interviews
- Patents
- Newsletters
- White Papers
- Book Chapters

- GreyNet 2004 survey (Boekhorst 2005)
  [www.greynet.org/greysourceindex/documenttypes.html](http://www.greynet.org/greysourceindex/documenttypes.html)
Grey Literature Defined

- “that which is produced on all levels of government, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers i.e. where publishing is not the main activity of the producing body”

- The Luxembourg Convention on Grey Literature, 1997; expanded New York, 2004

- [http://www.greynet.org/greynethome/listserv.html](http://www.greynet.org/greynethome/listserv.html)
Why is searching grey lit recommended?

- **Reduce bias**
  - Publication bias towards BIG results
  - Unpublished trial data inclusion can change statistical results

- **Currency**
  - Conference proceedings – significant time lag to journal publication for the ones that ever get published
  - Include emerging topics

- Include global research
Reasons for missing RCTs, CCTs in SRs

- Article(s) omitted or missing from a resource: 14%
- Insufficient or restricted search strategy: 14%
- Keywords or methodology not reported by author: 21%
- Published as abstracts, books, book reviews, brief reports, letters, proceedings or supplements, etc.: 21%
- Grey Literature: 21%
- Inadequate/inappropriate indexing: 67%

(Crumley 2005)

Used with permission: The Systematic Review Workshop: the Nuts and Bolts for Librarians. Health Sciences Library System-University of Pittsburgh. [http://www.hsls.pitt.edu/systematicreview] © Health Sciences Library System, University of Pittsburgh. All rights reserved.
Where is the grey literature?

- **Abstracts**: includes conference abstracts only
- **Unpublished**: trial registers, file-drawer data, individual trialists
- **Other**: reports, drug company, in press, letters, theses,

56% of abstracts will be published within 3 years. (Hopewell 2007)

Used with permission: The Systematic Review Workshop: the Nuts and Bolts for Librarians. Health Sciences Library System-University of Pittsburgh. [http://www.hsls.pitt.edu/systematicreview] © Health Sciences Library System, University of Pittsburgh. All rights reserved
Grey Literature: Trials, Regulatory Sources

- Clinical trial registries
  - ClinicalTrials.gov
  - WHO International Clinical Trials Registry Platform

- Regulatory data for drugs, devices
  - FDA
  - Health Canada Drug Product Database
  - European Medicines Agency
Identifying Conferences for inclusion

- Personal knowledge

- Identifying unknown conferences
  - Proceedings databases
    - OCLC Proceedings, OCLC Papers First
  - Bibliographic databases that include proceedings
    - Embase
    - Biosis Previews
    - Scopus
    - Web of Science
Selected sources for finding reports

- **US Federal Sources**
  - NTIS
    - www.ntis.gov/
  - AHRQ
    - www.ahrq.gov

- **Compilations**
  - HTA Database
    - www.crd.york.ac.uk/cms2web
  - Grey Lit Report NYAM
    - http://www.greylit.org/home
  - MedlinePlus: Organizational Web sites

- **Library Catalogs**
Choosing Filters
What is a search filter?

- A set of search terms designed to retrieve a specific kind of record
- Commonly to retrieve specific kinds of study designs
- But, can also retrieve sets of records for specific populations, topics, or anything that you might want to regularly systematically search
- Synonyms:
  - Clinical queries, hedges, highly sensitive search strategies, quality filters, evidence-based filters
Why use a search filter?

- Quickly and easily narrow down to the type of study design likely to produce relevant evidence

- Well designed, tested filters (validated filters) may give better results than informally developed search strategies

- Use of validated filters increases the perception that you are using a high quality search method overall
RCT filters

- Validated filters available, commonly used
- Cochrane provides them for Ovid MEDLINE, PubMed, and Embase
  - [http://handbook.cochrane.org/chapter_6/6_4_11_search_filters.htm](http://handbook.cochrane.org/chapter_6/6_4_11_search_filters.htm)

```
#1 randomized controlled trial [pt]
#2 controlled clinical trial [pt]
#3 randomized [tiab]
#4 placebo [tiab]
#5 drug therapy [sh]
#6 randomly [tiab]
#7 trial [tiab]
#8 groups [tiab]
#9 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8
#10 animals [mh] NOT humans [mh]
#11 #9 NOT #10
```

Highly Sensitive RCT filter for PubMed (2008 revision)
Where to Find Filters

- [https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/](https://sites.google.com/a/york.ac.uk/issg-search-filters-resource/)
  - InterTASC Information Specialists' Sub-Group (Centre for Reviews and Dissemination)

- [http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx](http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx)
  - Filters by the Hedges team (McMaster University)

  - NCBI (PubMed) clinical queries
Developing the Main Search
Who is doing what during main search development

**Investigators**
- Evaluating preliminary results, providing feedback
  - % on topic
  - Amount retrieved
  - Concepts to add, subtract
- Developing inclusion exclusion, data criteria and processes

**Librarian**
- Harvesting, testing search terms
- Documenting searches
- Explaining search to investigators
- Delivering preliminary results to investigators
- Modifying search as needed
- Getting peer review for search
Two Basic Lit Search Skills

- **Concept analysis, representation**
  - What word(s) best represent your concepts?
    - What synonyms bring in new on topic citations?
  - How many of the concepts in your question do you need to search to get the articles you need?

- **Syntax**
  - Meeting database’s expectations for input
  - Understanding the operators, modifiers to search allowed for each database
Building a concept module starting with MeSH

- **What is MeSH?**
  - Controlled Vocabulary
  - Indexers assign one word or phrase to represent a concept that can be said many ways

- **How is it used?**
  - Added it to records for articles on the topic

- **Example:** “Patient Compliance”[MeSH]
  - Used for: Patient Cooperation; Patient Adherence; Patient Non compliance; Patient Noncompliance; Patient nonadherence…..
MeSH Broader, Narrower Terms

Attitude to Health

Patient Acceptance of Health Care

Patient Compliance
  Medication Adherence

Patient Participation

Patient Satisfaction

Patient Preference

Treatment Refusal

Treatment. Perceived respect affects appointment adherence.

[No authors listed]
No Abstract

PMID: 23980292 [PubMed - indexed for MEDLINE]
Building a Concept Module starting from MeSH

**Process**

- Search for MeSH term that best matches your concept
  - Ex: "caregivers" [MeSH Terms]
- Find the Entry Terms for that MeSH term
  - Entry terms = synonyms
- Search for the Entry terms in title, abstract, and author keywords.
  - They will retrieve additional records not indexed to the MeSH term
  - Test to see if they bring in new relevant records
- Expand with other synonyms for your concept found in the literature
Caregivers

Persons who provide care to those who need supervision or assistance in illness or disability. They may provide the care in the home, in a hospital, or in an institution. Although caregivers include trained medical, nursing, and other health personnel, the concept also refers to parents, spouses, or other family members, friends, members of the clergy, teachers, social workers, fellow patients.

Year introduced: 1992

PubMed search builder options

MeSH term

Subheadings:
- classification
- economics
- education
- ethics
- history
- legislation and jurisprudence
- organization and administration
- psychology
- standards
- statistics and numerical data
- supply and distribution
- trends
- utilization

Restrict to MeSH Major Topic.
Do not include MeSH terms found below this term in the MeSH hierarchy.

Tree Number(s): M01.085, M01.526.485.200, N02.360.200
MeSH Unique ID: D017028

Entry Terms:
- Caregiver
- Carers
- Carer
- Care Givers
- Care Giver
- Spouse Caregivers
- Caregiver, Spouse
- Caregivers, Spouse
- Spouse Caregiver
- Family Caregivers
- Caregiver, Family
- Carers, Family
- Family Caregiver

All MeSH Categories
- Persons
- Caregivers

Terms I would test in title, abstract, author keywords:
- Caregiver
- Caregivers
- Carer
- Carers
- Care giver
- Care givers

All others shown will be retrieved with these words
Text Words

- **Words anywhere in the record**
  - Can specify where in the record the words appear
    - Title
    - Abstract
    - Author keywords
    - Author affiliation
    - etc

- **Keywords**
  - Sometimes used as synonym for textwords
  - Can also mean author supplied keywords
Testing Entry terms – PubMed example

<table>
<thead>
<tr>
<th>History</th>
<th>Add to builder</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#15 Add</td>
<td>Search (#14 NOT (#11 OR #12))</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>#14 Add</td>
<td>Search caregivers[ot]</td>
<td>840</td>
<td></td>
</tr>
<tr>
<td>#13 Add</td>
<td>Search (#12 NOT #11)</td>
<td>18606</td>
<td></td>
</tr>
<tr>
<td>#12 Add</td>
<td>Search caregivers[tiab]</td>
<td>30387</td>
<td></td>
</tr>
<tr>
<td>#11 Add</td>
<td>Search &quot;caregivers&quot;[MeSH Terms]</td>
<td>23529</td>
<td></td>
</tr>
</tbody>
</table>

Formula:
1. Search MeSH term (#11)
2. Search entry term in the title or abstract [tiab] (#12)
3. Isolate the new records found by the entry term (#13)
   1. Examine them – are enough about caregivers to include it?
4. Search entry term in the author keywords [OT]
5. Isolate the new records found by the entry term in comparison to the MeSH and other terms you are going to include (#15)
   1. Nothing new retrieved this time, don’t need to include in search
**Result of testing all caregiver terms**

Retrieval increased from 23,529 to 55,735

Next step: add other concept modules to narrow scope, reduce retrieval

<table>
<thead>
<tr>
<th>Search</th>
<th>Add to builder</th>
<th>Query</th>
<th>Items found</th>
</tr>
</thead>
<tbody>
<tr>
<td>#31</td>
<td>Add</td>
<td>Search ((#18 OR #16 OR #17 OR #21 OR #22 OR #27))</td>
<td>55735</td>
</tr>
<tr>
<td>#27</td>
<td>Add</td>
<td>Search (carer[tiab] OR carers[tiab])</td>
<td>9340</td>
</tr>
<tr>
<td>#22</td>
<td>Add</td>
<td>Search (&quot;care giver&quot;[ot] OR &quot;care givers&quot;[ot]))</td>
<td>18</td>
</tr>
<tr>
<td>#21</td>
<td>Add</td>
<td>Search (&quot;care giver&quot;[tiab] OR &quot;care givers&quot;[tiab])</td>
<td>2067</td>
</tr>
<tr>
<td>#17</td>
<td>Add</td>
<td>Search caregiver[ot]</td>
<td>666</td>
</tr>
<tr>
<td>#16</td>
<td>Add</td>
<td>Search caregiver[tiab]</td>
<td>16275</td>
</tr>
<tr>
<td>#12</td>
<td>Add</td>
<td>Search caregivers[tiab]</td>
<td>30387</td>
</tr>
<tr>
<td>#11</td>
<td>Add</td>
<td>Search &quot;caregivers&quot;[MeSH Terms]</td>
<td>23529</td>
</tr>
</tbody>
</table>
What if there isn’t MeSH for your concept?

- What did other systematic reviews do?
- Use Exemplar articles as a vocabulary source
- Do subject searches, find commonly used vocabulary
- Look in full text of key articles
Example: testing synonyms

**HCV term testing excerpt:**

1. Search hepatitis c  
   63029
2. Search hcv  
   37552
3. Search (hcv) NOT hepatitis c  
   2738

- Examine the retrieval in line 3
- **How many are about hepatitis c?**
- **More on or off topic?**
Building a modular search: combining synonyms

- “OR” together terms in a concept module
- Potential hepatitis c module:

hepatitis c OR "hepatitis non-a" OR "non-a, non-b hepatitis" OR hcv
Why go to such lengths?

- Be as inclusive as possible in your search
- Be sure you don’t include unneeded terms that add to reviewer burden.
Exercise 5

1. **Working in a group, read the citation and abstract on page 2.**

Then:

2. **Identify the 3 concepts in the systematic review search, and label each column with one of the concepts. (HINT: Think PICO)**

3. **Open the envelope. Arrange the words in the envelope in the columns so that synonyms and closely related concepts are in the same column.**
#3: Literature search methodology for a systematic review

Search Syntax Basics
Common Syntax Skills

- Using Boolean operators
- Using truncation
- Phrase searching
- Field Searching

Read the HELP section of each database – many variations exist!

Using PubMed

- PubMed Quick Start Guide
- Full Text Articles
- PubMed FAQs
- PubMed Tutorials
- New and Noteworthy
Boolean Basics

- 3 fundamental operators
  - AND,
  - OR,
  - NOT
AND

AND

=
AND to Connect Different Concepts

HIV infections AND adolescents AND treatment compliance

Results

- Less citations retrieved
- Those retrieved are more specific
#3: Literature search methodology for a systematic review

(AND

OR

) =
OR together synonyms

- HIV infections OR acquired immunodeficiency syndrome
- Adolescent OR teen

Results:
- More citations are retrieved
- More ways of saying the same thing are included
LI 1008.0: Systematic Reviews

#3: Literature search methodology for a systematic review
OR or AND?
ANDing all Darth’s stuff
ORing Darth’s stuff
What about NOT?
#3: Literature search methodology for a systematic review

- [ ] NOT
  - [ ]
  - [ ]
  - [ ]
  - [ ]

- [ ]
  - [ ]
  - [ ]
  - [ ]
## Primary use of NOT: comparing search results

<table>
<thead>
<tr>
<th>Rank</th>
<th>Search Term</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.</td>
<td>acquired immunodeficiency syndrome NOT hiv</td>
<td>38472</td>
</tr>
<tr>
<td>2.</td>
<td>acquired immunodeficiency syndrome</td>
<td>83741</td>
</tr>
<tr>
<td>1.</td>
<td>hiv</td>
<td>261134</td>
</tr>
</tbody>
</table>
Grouping the search concepts

( ) around OR sets

Example:
(HIV infections OR acquired immunodeficiency syndrome)
AND (Adolescent OR teen)
Truncation

- Get variations on a word without entering all of them
- EX: immun*
  - Retrieves immune, immunity, immunology, etc
- Read database help for symbols used, more options
Phrase Searching

- Requiring words to be together in the order specified
- Check database for syntax needed
  - Pubmed Example: “acquired immunodeficiency”
    - Finds phrases it recognizes from an index of phrases
    - Pretty good, but may not include the phrase you want
  - OVID MEDLINE Example: acquired immunodeficiency
    - Assumed phrase without any operator or quotes
    - With and without quotes gets same set
    - Exception: phrase with stop words in them.
      - EX: “labor and delivery”
Field Searching

- Asking for words to be in the specified field of the database.

- Example: phrase in title of article
  - PubMed Example: "HIV infections"[title]
  - OVID MEDLINE Example: hiv infections.ti.

- Commonly searched fields
  - Article title, abstract, author keywords, publication type
## Final Search Statement Example

### Appendix 1. PubMed search strategy, modified as appropriate for use in the other databases

<table>
<thead>
<tr>
<th>Search</th>
<th>PubMed</th>
</tr>
</thead>
<tbody>
<tr>
<td>#4 Search</td>
<td>((#1 AND #2) AND #3)</td>
</tr>
</tbody>
</table>

**Source:** Horvath 2012
Search quality assessment

- Did I accurately translate my question into search concepts?
- Did I use the right Boolean and proximity operators?
- Are there spelling errors?
- Did I use the correct line numbers in combining search statements?
- Is the translation of the MEDLINE search for other databases accurate?
- Did I include the relevant subject headings?

(Sampson and McGowan, 2006)
Other Search Skills needed for SRs

- Search translation between databases
- Deduplication of results
- Meticulous documentation
Standards for search documentation

- **For internal and external audiences**
  - Internal: Copy, paste search history from database to file
    - Include all search statements, number of hits
  - External: 8 commonly used elements (Mullins et al.)

- **Good documentation advice available**
  - Institute for Quality and Efficiency in Health Care 2015
  - Rader et al 2014
References


References cont.
