SEARCH FILTERS

Prepared by:

Cynthia Fraser, Information Specialist, Health Services Research Unit, Aberdeen
Jenny Harbour, Health Information Scientist, Healthcare Improvement Scotland
WEBEX GOOD PRACTICE

• Webex interface:
  • If you have a question use the chat box or the “raise hand” button
  • If you’re having technical difficulties message the host via the Webex chat function

• Tele-conference:
  • Remember to introduce yourself before speaking
  • Remember to speak into the telephone speaker
  • Don’t have side conversations with colleagues
  • Mute your phone if you are in a noisy environment
WHO ARE WE?

Cynthia Fraser
Information Specialist
Health Services Research Unit

Jenny Harbour
Health Information Scientist
Healthcare Improvement Scotland
• Search and synthesis element of Knowledge Into Action programme

• Need to develop skills within K&IS to support search and synthesis

• Expressed desire for filters support/training

• Part of an ongoing programme of seminars to support K2A implementation
OUTLINE

• Introduction
  – What are filters?
  – Types of filter
  – Performance measures
• Developing and testing search filters
• Finding and selecting a filter
• Appraising filters
• Translating filters
• Pros and cons of search filters
WHAT ARE SEARCH FILTERS?

- Search terms/strategies to identify a specific topic or aspect
- Intended for repeated use
- Routinely combined with subject specific terms

### Cochrane HSSS: sensitivity maximising

1. randomized controlled trial.pt.
2. controlled clinical trial.pt.
3. randomized.ab.
4. placebo.ab.
5. drug therapy.fs.
6. randomly.ab.
7. trial.ab.
8. groups.ab.
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8
10. exp animals/ not humans.sh.
11. 9 not 10
TYPES OF SEARCH FILTER

- Methodological filters
- Subject/focus filters
- Sensitive, precise or optimised/balanced
- Development method of filter
  - Pragmatic
  - Research based/objective
PERFORMANCE MEASURES

• Sensitivity:
  number of relevant records retrieved as a proportion of the total number of relevant records

• Precision:
  number of relevant records retrieved as a proportion of the total number of records retrieved

• Specificity:
  number of irrelevant records NOT retrieved as a proportion of the total number of irrelevant records

• Other measures include accuracy, number needed to read (NNR), etc
DEVELOPING AND TESTING FILTERS

• Objectively derived – gold/reference standards
• Identify candidate terms
• Select terms by testing performance
• Combine terms
• Test filter performance
GOLD OR REFERENCE STANDARD (GS)

• Set of studies that meet purpose of filter e.g. for RCT filter, identify a set of articles reporting RCTs

• Identify corresponding database records in MEDLINE – combine UIs with OR

• Ideally sample set should be representative of RCTs in MEDLINE – time period, subject content, terminology
Hand search a set of journals:

- Scan complete contents for reports of studies
- Variety of journals
- Include different time periods
- Labour intensive method

Hand search GS example

**Journals:**
BMJ, Controlled Clinical Trials, Contemporary Clinical Trials

**Years:** 2000, 2010

N=150
Included studies from systematic review(s):

- Original search for review had no ‘filter’
- GS will only be as comprehensive as search
- GS may be incomplete because of review inclusion criteria
- Multiple reviews - vary subject coverage

Included studies GS example

Included studies from 3 reviews on different topics
Reports indexed in MEDLINE

N=150
IDENTIFYING CANDIDATE TERMS

• Identify search terms that might be useful to include in filter
• Scan GS records for indicative MeSH and free text words or phrases
• Consult database thesaurus
• Expert opinion
• Frequency a guide but not stop-words, irrelevant terms etc
• Text analysis software
• Attempt to be objective as possible
SELECTING TERMS (1)

- Test performance to inform whether to include in filter
- How many records retrieved by terms
  - in total set
  - in gold standard

<table>
<thead>
<tr>
<th>TOTAL SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (bmj or “controlled clinical trials” or “contemporary clinical trials”).jn and (2000 or 2010) or 4500</td>
</tr>
<tr>
<td>1. Original subject search 4500</td>
</tr>
<tr>
<td>GS</td>
</tr>
<tr>
<td>2. 1.ui or 2.ui or 3.ui……….150.ui 150</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TEST TERMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. 1 and randomised.tw 200</td>
</tr>
<tr>
<td>4. 1 and randomized.tw 50</td>
</tr>
<tr>
<td>5. 2 and randomised.tw 100</td>
</tr>
<tr>
<td>6. 2 and randomized.tw 30</td>
</tr>
</tbody>
</table>
SELECT TERMS (2)

- Sensitivity = proportion of GS retrieved

<table>
<thead>
<tr>
<th>Term</th>
<th>GS retrieved</th>
<th>Total GS</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>randomised</td>
<td>100</td>
<td>150</td>
<td>(100/150) x 100 = 66.7%</td>
</tr>
<tr>
<td>randomized</td>
<td>30</td>
<td>150</td>
<td>(30/150) x 100 = 20.0%</td>
</tr>
</tbody>
</table>

- Precision = proportion of retrieved that are GS

<table>
<thead>
<tr>
<th>Term</th>
<th>GS retrieved</th>
<th>Total retrieved</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>randomised</td>
<td>100</td>
<td>200</td>
<td>(100/200) x 100 = 50.0%</td>
</tr>
<tr>
<td>randomized</td>
<td>30</td>
<td>50</td>
<td>(30/50) x 100 = 60.0%</td>
</tr>
</tbody>
</table>
## COMBINE TERMS

<table>
<thead>
<tr>
<th>Terms</th>
<th>Sensitivity</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomised.tw</td>
<td>66.7</td>
<td>50.0</td>
</tr>
<tr>
<td>Randomized.tw</td>
<td>20.0</td>
<td>60.0</td>
</tr>
<tr>
<td>Random$.tw</td>
<td>80.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Cross-over.tw</td>
<td>5.0</td>
<td>95.0</td>
</tr>
<tr>
<td>Randomized controlled trial.pt</td>
<td>95.0</td>
<td>10.0</td>
</tr>
</tbody>
</table>

- Most sensitive – randomized controlled trials.pt or random$.tw
- Most precise – cross-over.tw or randomized.tw or randomised.tw
- Cut-off e.g. s ≥50.0 + p ≥20.0 – random$.tw (or randomised.tw)
- Statistical analysis
## TESTING FILTER PERFORMANCE AGAINST GS (INTERNAL)

<table>
<thead>
<tr>
<th>Terms</th>
<th>Total set retrieved</th>
<th>GS retrieved</th>
<th>Sensitivity</th>
<th>Precision</th>
<th>Strategy</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most sensitive</td>
<td>3500</td>
<td>150</td>
<td>100.0</td>
<td>5.0</td>
<td>Sensitivity maximising</td>
<td>Comprehensive search</td>
</tr>
<tr>
<td>S\geq50+P\geq20</td>
<td>900</td>
<td>135</td>
<td>90.0</td>
<td>15.0</td>
<td>optimised</td>
<td>pragmatic</td>
</tr>
<tr>
<td>Most precise</td>
<td>450</td>
<td>90</td>
<td>60.0</td>
<td>20.0</td>
<td>Precision maximising</td>
<td>Scoping search</td>
</tr>
</tbody>
</table>
TESTING IN THE REAL WORLD (EXTERNAL)

• Filter development often undertaken in small, selective sample of records
• External validation is a more robust assessment
• Use different gold standard
  – some primary studies that develop filters do this
  – reviews comparing filter performance using common GS
• Testing in complete database
HOW FILTERS ARE TESTED: SUMMARY

**DEVELOPMENT STUDY**

1. Derive filter
2. Test filter in derivation set
3. Test performance in validation set
4. Gold standard A
5. Derivation set
6. Validation set

**PERFORMANCE COMPARISON**

1. Test all filters in gold standard B
2. Developed filter
3. Published filters
4. Gold standard B
FINDING AND SELECTING A FILTER (1)

• **ISSG website**
  – collate reports of methods filter development and performance
  – methods of some are appraised

• Large research projects
  – **Cochrane** RCT
  – **Hedges team** (Ovid)

• Organisations with extensive experience
  – **SIGN**
  – Centre for Review & Dissemination (CRD)
### Evaluations of the performance of filters can be found below the table

<table>
<thead>
<tr>
<th>Database</th>
<th>Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>CINAHL</td>
<td>SIGN strategy [undated] [Ovid]</td>
</tr>
<tr>
<td>EMBASE</td>
<td>SIGN strategy [undated] [Ovid]</td>
</tr>
<tr>
<td></td>
<td>BMJ Clinical Evidence strategy [undated] [Ovid]</td>
</tr>
<tr>
<td></td>
<td>SIGN strategy [undated] [Ovid]</td>
</tr>
<tr>
<td></td>
<td>de Freitas AE, Herbert RD, Lattimer J, Farreira PH. Searching the LILACS database for Portuguese- and Spanish-language randomized trials in physiotherapy was difficult. Journal of Clinical Epidemiology 2006;59(3):233-7. [RCTs in physiotherapy]</td>
</tr>
<tr>
<td></td>
<td>Elsinga A, Sloepbergen N, Clarke M. The sensitivity and precision of search terms in phases I, II and III of the Cochrane High Risk Sensitivity Search Strategy for identifying reports of randomized trials in</td>
</tr>
</tbody>
</table>
6.4.11 Search filters

6.4.11.1 The Cochrane Highly Sensitive Search Strategies for identifying randomized trials in MEDLINE

Box 6.4.a: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2003 revision); PubMed format
Box 6.4.b: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); PubMed format
Box 6.4.c: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity-maximizing version (2008 revision); Ovid format
Box 6.4.d: Cochrane Highly Sensitive Search Strategy for identifying randomized trials in MEDLINE: sensitivity- and precision-maximizing version (2008 revision); Ovid format

6.4.11.2 Search filters for identifying randomized trials in EMBASE
FINDING AND SELECTING A FILTER (4)
FINDING AND SELECTING A FILTER (5)

- Randomised Controlled Trials
- Observational Studies
- Diagnostic Studies
- Economic Studies
- Patient issues

SYSTEMATIC REVIEWS

The search filter used by SIGN to retrieve systematic reviews is an adaptation of the systematic reviews filter designed by the Health Information Research Unit of the McMaster University, Ontario. The systematic reviews filter emphasises specificity rather than sensitivity.

Medline

1. Meta-Analysis as Topic/  
2. meta analy$ tw.  
3. metaanaly$ tw.  
4. Meta-Analy$  
5. (systematic adj(review$ or overview$)).tw.  
6. exp Review Literature as Topic/  
7. or1-8  
8. cochrane ab.  
9. embase ab.  
10. (psychlit or psychlit).ab.  
11. (psychlit or psychlit).ab.  
12. (cinhit or cinhit) ab.  
13. science citation index ab.  
14. bids.ab.  
15. cancerlit.ab.  
16. or15-18  
17. reference list$ ab.  
18. bibliography ab.  
19. hand-search$ ab.  
20. relevant journals ab.  
21. manual search$ ab.  
22. err17-21  
23. selection criteria ab.  
24. data extraction ab.  
25. 23 or 24  
26. Review/  
27. 25 and 26  
28. Comment/  
29. Letter/  
30. Editorial/  
31. animal/  
32. human/  
33. 31 not (31 and 32)  
34. err28-30,33  
35. 7 or 16 or 22 or 27  
36. 35 not 34
## FINDING AND SELECTING A FILTER (2)

<table>
<thead>
<tr>
<th>Type of question</th>
<th>Example question</th>
<th>Type of filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical effectiveness</td>
<td>Is venlafaxine effective for the treatment of menopausal sweats?</td>
<td>Systematic review, guidelines, RCT, prognosis</td>
</tr>
<tr>
<td>Cost effectiveness/disinvestment</td>
<td>Is treating obese patients with drugs in primary care cost effective?</td>
<td>Economic evaluation, systematic review, guidelines, quality of life</td>
</tr>
<tr>
<td>Test selection</td>
<td>Which imaging technique is most effective for detecting chondral lesions in the hip joint?</td>
<td>Diagnostic test accuracy, RCT</td>
</tr>
<tr>
<td>Patient safety</td>
<td>What are the risks to patient safety of immunisation in children?</td>
<td>Adverse events, systematic review, guidelines, RCT</td>
</tr>
<tr>
<td>Population/public health</td>
<td>Which ethnic groups are at greatest risk of developing cardiovascular disease?</td>
<td>Epidemiology, aetiology, observational studies</td>
</tr>
<tr>
<td>Patient experience</td>
<td>What are patient’s views on awake craniotomy?</td>
<td>Qualitative studies, observational studies, quality of life</td>
</tr>
</tbody>
</table>
FINDING AND SELECTING FILTERS (3)


• Glanville JM, Lefebvre C, Miles JNV, Camosso-Stefinovic J. How to identify randomized controlled trials in MEDLINE: Ten years on. JMLA. 2006;94(2):130-6.

APPRAISING FILTERS

- Things to consider when appraising a filter/filter paper:
  - Does the purpose of the filter align with your needs?
    - Performance measures, database & interface
  - Is it clear how the filter was developed?
  - Are the methods used to develop the filter appropriate/ fit for purpose?
  - Did the filter perform at a suitable level when tested?
  - Has the filter been compared with other filters?
  - Are there any other limits or errors to consider?
TRANSLATING FILTERS

• Why translate filters?
  – Published filters applicable to specific database and platform
  – Translate filter for use in a different platform e.g. Ovid Medline vs. Pubmed
  – Translate filter for use in a different database e.g. Medline filter translated for Embase
  – Translate a filter for use in a different database and different platform e.g. Ovid Medline filter for EBSCOHost Cinahl
TRANSLATING FILTERS

• Considerations when translating filters:
  – Thesaurus term mapping
  – Exploding subject headings e.g. MeSH
    • Automatic (Pubmed) vs. manual (Ovid)
  – Field codes e.g.
    • .tw vs. :ti,ab
    • ‘/’ vs. mh
  – Truncation e.g. ‘$’ vs. ‘*’
  – Proximity e.g. ‘adjn’ vs. ‘near’ vs. not available
  – Limit options e.g. language, human, dates
SEARCH FILTERS – PROS AND CONS

Pros

• Save time in development of search strategy
• Reduce number of abstracts to be screened
• Pragmatic solution when time and resources limited

Cons

• Sensitive filters not achieve 100% (but what search does?)
• Filter performance tends to be poorer in ‘real word’
• Translation problems
• Difficult to choose which to use
RESOURCES

• ISSG Search Filters Resources
  – http://www.york.ac.uk/inst/crd/intertasc/

• McMaster University Hedges Team
  – http://hiru.mcmaster.ca/hiru/HIRU_Hedges_home.aspx

• Cochrane Handbook
  – http://www.cochrane-handbook.org/

• SIGN search filters
  – http://www.sign.ac.uk/methodology/filters.html


• Training course: ”Getting the best out of search filters”. More information at: http://php.york.ac.uk/inst/yhec/web/training/index.php
CONTACTS

Cynthia Fraser
Information Specialist
Health Services Research Unit, Aberdeen
E-mail: c.fraser@aberdeen.ac.uk

Jenny Harbour
Health Information Scientist
Healthcare Improvement Scotland
E-mail: jenny.harbour@nhs.net