Getting better all the time: Considerations and approaches for LSR searching

Robin Featherstone
*Information Specialist, Alberta Research Centre for Health Evidence (ARCHE), Cochrane Child Health, Canada*

Steve McDonald
*Information Specialist, Cochrane Australia, Monash University, Australia*
Webinar Outline

• Part 1 – General considerations
• Part 2 – Cochrane LSR approach
• Q&A
Disclosures

• We are both members of the Living Evidence Network (LEN)

• We are both involved in developing and/or evaluating LSR search methods and tools

• Steve McDonald is the co-lead of the LEN Search Interest Group
Agenda – Part 1

• Introduction to LSR information retrieval
• Automated and facilitated approaches
• Assessment of search performance
• Maintenance of search strategies
Search Implications of LSR approach

We define an LSR as a systematic review that is continually updated, incorporating relevant new evidence as it becomes available. In practice, this means continual surveillance for new research evidence through ongoing or frequent searches and the inclusion of relevant new information into the review in a timely manner so that the findings of the systematic review remain current.

Elliott et al. 2014 *PLoS Med*
Features of Cochrane LSR Approach

• Applies to any review type (e.g. RCTs, qualitative)
• Retains core systematic review methods
• Protocol pre-specifies LSR-specific methods
• Includes explicit and a priori commitments to frequent search and updating
• Starts with a standard ‘baseline’ review
Challenges for Information Specialists (IS)

- Addressing workload increases
- Translating bespoke search strategies into LSR production models
- Accounting for potential publication bias
- Assessing and revising strategies
- Maintaining search performance over time
IS Searching Goal

Maximize efficiency while ensuring quality
Adapting Bespoke Search Strategies

Considerations

• How well did the original search perform?
• Can it be replicated?
• Can it be automated (in full or in part)?
• Could a highly precise LSR search supplement the original?

What revisions to the search are needed:

• To reduce workload?
• To ensure precision?
Search Reporting – Why Standards Matter

Table 1. Literature search strategy

<table>
<thead>
<tr>
<th>Topic</th>
<th>Search terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Food environment</td>
<td>Food environment OR nutrition environment OR retail food OR neighbourhood OR neighborhood OR environment OR food desert* OR food swamp OR food availability OR food cost OR food affordability OR food pric* OR food quality.</td>
</tr>
<tr>
<td>Retail food outlets</td>
<td>Supermarket OR grocery store OR convenience store OR corner store OR dollar store OR fast food OR restaurant OR food store OR bodega OR tienda.</td>
</tr>
<tr>
<td>Dietary intake</td>
<td>Food OR fruit OR vegetable OR diet* OR nutrition OR processed food.</td>
</tr>
<tr>
<td>Weight- and health-related outcomes</td>
<td>Obes* OR overweight OR BMI OR body mass index OR waist circumference OR anthropometric OR health OR cardiovascular OR cancer OR diabetes OR hypertension OR disease OR illness.</td>
</tr>
<tr>
<td>Socio-economic status</td>
<td>Income OR disparity OR equity OR inequity OR inequality OR disadvantage OR poverty OR depriv* OR marginaliz*.</td>
</tr>
<tr>
<td>Canada</td>
<td>Canada OR Canadian OR British Columbia OR Alberta OR Saskatchewan OR Manitoba OR Ontario OR Quebec OR Nova Scotia OR New Brunswick OR Prince Edward Island OR Newfoundland OR Yukon OR Northwest Territories OR Nunavut.</td>
</tr>
</tbody>
</table>

* A Boolean search function indicating truncation, allowing multiple forms of a given word (e.g., depriv* identifies deprived, deprivation).

- What are the reporting problems?
- Can we replicate?
- Can we adapt into an LSR search?
Automated Approaches - Alerts

Considerations

- Do search sources (databases, registers, etc.) support auto-alerts?
- Will auto-alerts match planned update frequency?
- Is it more efficient to use auto-alerts or to run saved searches?
- How to apply date limits?
- How to identify and remove duplicates?
- Who will manage results?
Objectives of LSR Auto-alerts

Retrieve precise, ready-to-screen, unique search results on predictable intervals
Accounting for Publication Bias - Facilitated Approaches

Complementary search methods

- Cited/citing references
- Handsearching
- Contacting study authors

Grey literature sources

- Trial registers
- Agency reports
- Industry websites
- Funded research databases
- Etc., …
Facilitated Searches in LSR Production Models

Considerations

- Which complementary search methods or grey literature sources are likely to yield new unique evidence?
- How to integrate results with auto-alerts?
- Should frequency be the same as auto-alerts (in full or in part)?
- Who should conduct?
Objectives of LSR Facilitated Searching

Complement (not replicate) auto-alerts

Created by Laymik from Noun Project

Created by Yazmin Alanis from Noun Project
Potential Pitfalls of LSR Searching

"I never make the same mistake twice. I strive for five or six times."

©Gabriel (UTS) 06-09-07
Assessing LSR Searches

Considerations

• Which assessment methods to use?
  ➢ Sensitivity, precision, number needed to read (NNTR), accuracy, yield? (Cooper, 2018)

• How to maximize knowledge gained from the review process?
  ➢ Utility of included studies from the ‘baseline’ review
## Case Study: Retrospective Search Assessment

- Data to inform a 2018 update of a 2016 systematic review
- Number of included studies = 139

<table>
<thead>
<tr>
<th>Database</th>
<th>Records retrieved</th>
<th>Total included studies retrieved</th>
<th>Precision</th>
<th>NNTR (Number Needed to Read)</th>
<th>Unique studies retrieved</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEDLINE</td>
<td>4037</td>
<td>118</td>
<td>0.029</td>
<td>34</td>
<td>3</td>
</tr>
<tr>
<td>Embase</td>
<td>2623</td>
<td>121</td>
<td>0.046</td>
<td>22</td>
<td>6</td>
</tr>
<tr>
<td>CENTRAL</td>
<td>970</td>
<td>128</td>
<td>0.132</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>CINAHL</td>
<td>254</td>
<td>34</td>
<td>0.134</td>
<td>7</td>
<td>0</td>
</tr>
</tbody>
</table>
Value of Search Assessment

- Determine time needed to screen results from each source (and combinations of sources)
- Identify sources that contribute unique studies
- Target revisions to improve search precision
Maintenance Required

Considerations

- Is the review question maturing over time? Are search terms evolving?
- Which new index terms (e.g., MeSH) should be integrated into the strategy?
- Will database enhancements impact auto-alerts?
- Are new evidence sources available?
- How often should the strategy be examined?
Objectives of LSR Search Assessment and Maintenance

Improve performance and enhance strategies on a recurring basis

“It’s getting better all the time …”
Summary and Recommendations

• Anticipate changes to the search methods
• Pre-specify revisions, assessment and maintenance activities in the protocol
• Combine automated and facilitated strategies to reduce workload and account for potential publication bias
• Maximize opportunities to improve search performance
Agenda – Part 2

- Features and requirements of Cochrane LSRs
- Enablers to support LSRs
- Cochrane LSR pilots
Features of Cochrane LSRs Search

- Full versus partial
  - Principle of full searches
- Compliance with mandatory MECIR standards
  - Database searches (C24)
  - Trial Registers (C27)
  - Reference checking (C30)
- Search methods text in protocol/review
- ‘What’s new’ and History sections indicate current status of search results in the review
- Annual review of search methods
Enablers for Study Selection

Involves taking research curation outside the confines of individual reviews

Uses three core ‘technologies’:

1. Human effort in Cochrane Crowd
2. Machine Learning
3. [Re-use of data]
Cochrane Crowd

You can make a difference!

Become a Cochrane citizen scientist. Anyone can join our collaborative volunteer effort to help categorise and summarise healthcare evidence so that we can make better healthcare decisions.

A platform for crowdsourced **micro-tasks** that helps produce high quality health evidence
Results to Date

- 75,000 trials found
- 2 million classifications
- 9000 contributors
- 99% accuracy
**Text Mining**

Deriving high-quality information from **text**

**Machine Learning**

Models that **learn from data** to make predictions or decisions
Machine Learning Classifiers

Models that **learn from data** to make predictions or decisions
RCT Classifier

- Available through CRS Web and EPPI-Reviewer
  - trained on 400,000 classifications by Cochrane Crowd
- Assigns a probability score (0-100) to each citation
- Recall of 99.8% at 10% (0.1) threshold
- Can reduce screening load by 60-80% by discarding the ‘very unlikely’ to be RCT citations
- Crowd may help in screening citations (score 11-100)

- 25,000 records
- 15,655 very unlikely to be RCT
- 99.9% correct
‘Screen 4 Me’ Workflow

Start: conduct usual review searches

- Are these records already known NOT to be RCTs? [Yes/No]
  - Yes → Existing data
  - No → Are these records very unlikely to be RCTs? [Yes/No]
    - Yes → Are these records RCTs according to Cochrane Crowd? [Yes/No]
      - Yes → End (Manual screening of remainder)
      - No → End (Manual screening of remainder)
    - No → End (Manual screening of remainder)
Cochrane LSR Pilots

- Ongoing evaluation to explore feasibility, and implications for contributors, of LSR processes and workflows
- Participation of author teams, editorial staff, information specialists
- Five Cochrane LSRs published as of July 2018:
  1. Interventions for increasing fruit and vegetable consumption in children aged five years and under
  2. Delayed antibiotic prescriptions for respiratory infections
  3. Three reviews on the prevention and treatment of venous thromboembolism (VTE) in patients with cancer. (To inform guidelines by the American Society of Hematology)
• Monthly searches of Cochrane, MEDLINE and Embase, and trial registers

• 2,600 citations since Sep 2017 sent to RCT Classifier
  • Citations scoring 10-99 (1260; ~50%) >>>> AUTHORS
  • Citations scoring 0-9 (1340) >>>> CROWD

• 16 new RCTs; 1 RCT from among citations sent to Crowd
Documenting and Reporting the Search

- For each new publication version
  - amend ‘Results of the search’ section
  - update PRISMA flowchart

Results of the search

We ran searches for the previous reviews (Wolfenden 2012; Hodder 2017; Hodder 2018) and this review update, which together generated a total of 25,480 citations (24,661 previous reviews; 819 this review update). Screening of titles and abstracts for the review update identified 91 records (737 in total, including 646 from the previous reviews) for formal inclusion or exclusion.
Cochrane Database of Systematic Reviews
17 MAY 2018 DOI: 10.1002/14651858.CD008552.pub5
• Monthly update of ‘What’s new’ section

What's new

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>26 June</td>
<td>Amended</td>
<td>This is a Living Systematic Review. Searches are run and screened monthly. Search results up to 25 January 2018 are included in the current update (published May 2018). In addition, the team continues with the monthly screening (last search date 25 May 2018) and has found 4 new studies and 1 new ongoing study that will be included in a future update</td>
</tr>
</tbody>
</table>
Final Words

• Don’t panic
• Anticipate workload
• Start small
• Be meticulous – organization pays off
• Assess, adapt and improve

• IS expertise required
Questions

Created by Gilbert Bages
from Noun Project
Acknowledgements

• Anneliese Synnot and Tari Turner for organizing and promoting this webinar
• James Thomas and Anna Noel-Storr, as members of Project Transform, for their contribution to the slides on machine classifiers and Crowd
• Cochrane Information Specialists who shared their LSR experiences with us
• Justin Clark for his review of the slideshow presentation
References & Recommended Resources


Cochrane Community. Living Systematic Reviews. 2018. Available from: http://community.cochrane.org/review-production/production-resources/living-systematic-reviews


References & Resources, cont.

CRS Web videos and Quick References Guides:

Sending Records to the Classifier:
Video: https://youtu.be/IOoZM_KbjvQ
PDF Quick Ref Guide:

RCT Classifier example for Living Systematic Reviews:
Video: https://youtu.be/F_67pRaP5TI
PDF Quick Ref Guide: