

Report of Cochrane Canada living systematic review workshop

13 & 14 May 2017

David Braley Health Sciences Centre
McMaster University, Hamilton, Canada

Prepared by Anneliese Synnot and Julian Elliott,
with input from workshop attendees.



Contents

Background	3
Aims	3
Speakers, facilitators and attendees	3
Workshop structure	3
Themes and concepts discussed	4
Defining a living systematic review	4
Minimum criteria or attributes for LSRs	4
Rationale for a living systematic review	5
Searching and LSRs	5
Production and publication of LSRs	6
Utilising technology and citizen science	6
Methods for updating meta-analysis	7
Intersections with living guidelines	8
Policy and implementation implications	8
Opportunities for consumer engagement	9
Concerns or caveats with engaging consumers in LSRs	9
Evaluating the process and impact of LSRs	9
Towards a research agenda	10
Appendix: Organising committee, speakers, facilitators and attendees	12

Background

In May 2017, the teams from Cochrane Canada and Cochrane Australia co-hosted a two-day Living systematic review (LSR) workshop immediately following the Cochrane Canada Symposium, in Hamilton, Canada. This report summarises the presentations and discussions held, and proposes a number of recommendations for next steps and future activities.

Aims

The aims of the workshop were to:

1. Introduce attendees to the state-of-the-science in LSRs, providing a practical understanding of LSR methods, processes, publication implications and technological enablers;
2. Explore in more detail the opportunities, challenges, methods and processes for producing, publishing, and evaluating LSRs, and to begin to shape a research agenda.

Speakers, facilitators and attendees

Speakers and facilitators included people with experience piloting LSRs, and those with expertise in systematic review methods, including searching and statistics, guidelines, citizen science, and knowledge translation (see Appendix 1).

Attendees included members of the [LSR Network](#), people who were exploring various components of LSRs and people who were new to LSRs (see Appendix 1). It included researchers, systematic review producers, people from health technology assessment agencies and guideline developers. Approximately 40 people attended day one and 20 attended day two. Attendees came from many different countries, and from within and outside of the LSR Network.

Workshop structure

Day one of the workshop was open to anyone interested in LSRs and included a series of rapid presentations on:

- introducing the living systematic review concept;
- presenting the state-of-the-science in methods, production and publication;
- exploring the opportunities to extend LSRs into policy and practice, including living guidelines;
- sharing the experiences of six teams piloting or exploring LSRs, in a variety of topic areas (see slide sets for more information).

The slide sets for most of these presentations are available [online](#).

Day two of the workshop included primarily small and large group discussions and was invitation only.

Themes and concepts discussed

Over the two days participants provided valuable experiences and insights about the practical and methods considerations of LSRs, and shared ideas for how to move the research agenda forward. Rapporteurs took notes for all large and small group discussions on both days. Discussion was aimed at generating ideas and exploring concepts rather than seeking consensus about how to move forwards.

Below we highlight important points that came out of the discussions. Additionally, some of the ideas generated were included in the 4-part series on LSRs that was subsequently published in the [November 2017 issue](#) of the Journal of Clinical Epidemiology.

Defining a living systematic review

The following definition was taken into the workshop: “A systematic review that is continually updated, incorporating new evidence as it becomes available.”

There was some discussion of whether a living systematic review is a *process* or a *product*, and enthusiasm to revisit the definition to better incorporate the fact that an LSR involves a proactive process with the ultimate aim of retaining currency for end users, however they might define currency. Systematic review updating could be thought of as a spectrum, LSRs are optimising the process for the most frequent updating.

At workshop end, the following alternative definition was proposed (but no consensus reached): “LSR is a *process* [OR: LSR is a *systematic review* that follows a proactive process] that enables continuous surveillance for new evidence for incorporation into systematic reviews, so that the findings of the systematic review remain current, in order to meet identified needs of end users.”

Minimum criteria or attributes for LSRs

Given that the definition of a living systematic review is broad, and we are increasingly seeing researchers undertaking LSRs, there was support for devising minimum criteria for LSRs. These would help to operationalise the concept, and ensure consistency across LSRs. It was envisaged that these minimum criteria could present the first step towards reporting guidelines for LSRs or a PRISMA extension.

Suggestions for minimum criteria for LSRs included:

- Conclusions must be current or up to date (although defining ‘current’ or ‘up to date’ should ideally be by users/stakeholders)
- There is continual evidence surveillance underpinning the conclusions
- The evidence surveillance is ‘proactive’ but the trigger for adding new studies should be when they affect the certainty of the evidence (as opposed to automatically including any new study, or data/information about a new study, even if it makes no difference on the review results)
- An explicit commitment to maintain the LSR over time
- A priori specified process (or methods) for the living approach

- Core systematic review methods remain the same (i.e. retains validity of a standard systematic review)
- LSR is a systematic review “with a certain streamlined process and method”
- Must be a credible systematic review (i.e. conducted to a high standard)

Rationale for a living systematic review

There was a realisation that the thinking to date about why we should explore LSRs has been done from the perspective of producers (i.e. focus on streamlining workload) working on an assumption that review users would find continually up to date evidence beneficial. While we have no reason to think this is not the case, we don't yet have a good idea of what potential users of LSRs (including guideline developers, consumers, health professionals, policymakers etc) think about the concept, whether they would find it particularly appealing, and how they would like to see the concept operationalised, or the potentially evolving and changing findings communicated.

Searching and LSRs

What are the implications of running (and revising) searches over time?

- The conditions when searches should be revised could include: new MeSH terms, errors spotted, databases changed, or a change in PICO criteria. It is likely that new librarians or search specialists will have different preferences for how searches are designed.
- There is also an opportunity to improve search precision through regular maintenance of search strategies
- If searches are revised, do we need to go back and run all previous searches? And does this differ by reason for the revised search (i.e. if PICO changes versus new MeSH term?)
- How are these search changes peer reviewed? Does it need to go in an updated protocol?
- Workload is likely to be significantly greater for search specialists (or those managing search alerts) as compared to typical SR commitment related to searching.

How can duplicates be managed?

- By re-running searches frequently, there are likely to be considerable duplicates at each search re-run (i.e. duplicate between sources, multiple times in the same sources; duplicates between update and previous update), which would be time-consuming for information specialists.
- Could there be server site deduplication, or could new programs (i.e. Covidence) help?
- Does the plan for managing duplicates and “associated publications” (e.g. the same study in different sources) need to be addressed in the protocol?

How should search update frequency be determined?

- The preferred frequency for updating electronic searches might be topic-dependent, and related to the rate of new studies being published

- With a living model, it would be ideal to be able to track registered trials once identified to be able to include them as soon as they are published
- How many person hours are needed each time the searches are re-run to collate/organise citations for screening?

Is there an opportunity for a flexible model, using abridged searches for LSRs?

- Could an LSR search be more flexible (and responsive to what is found and where)? Potentially some sources may not need frequent searching (i.e. grey literature)?
- Could a new model be used, whereby abridged surveillance searches are run, and the full search is only run when new evidence (that potentially affects the results in some important way) is found via this horizon scanning?

Production and publication of LSRs

What might be lost by moving from a traditional to a continuous update model?

- The current infrequent update model facilitates something of a 'stocktake' of the review, as the review scope, methods and results are reviewed with fresh eyes by (often new) authors, editors and peer reviewers.
- Given the continuous nature of updating an LSR, these review 'stocktakes' will need to be scheduled in, but at what frequency? And to what degree would we want consistency in the authors, editors and peer reviewers over time?

How could we incentivise a versioning publication system for journals and authors?

- LSRs should ideally be published in academic journals, as this maximises findability, allows a clear record of changes over time, ensures appropriate peer review and feeds academic credits.
- But neither journals nor authors would want reviews re-published with a new citation every time new evidence, data or information is added. As such, a versioning system, in which new evidence can be added to the original manuscript, without triggering a new publication, is warranted. This is a model that some journals are exploring but is not a widespread system and may be difficult to incentivise journals to do.

Utilising technology and citizen science

Technology-related possibilities are evolving quickly, but connectivity between systems is challenging.

- The Centralized search service project is including an increasing number of databases, such that one day the vision is that we won't need to search the main databases (MEDLINE, EMBASE etc) individually.
- Machine classifiers are currently working very well with RCTs, and increasingly being developed to be able to identify other study designs, and could allow repositories of other study designs in CENTRAL

- Would be interesting to know if handsearching can be automated (e.g. being able to take reference lists of included studies to identify unpublished studies or conference abstracts).
- At the moment, the key is to identify the best tool for each task and to ensure long-term sustainability

Concerns about ‘validity’ of Crowd decisions on RCT study identification are unfounded

- There was some scepticism about the quality of screening conducted by Cochrane Crowd. To ensure accuracy, all records are reviewed and agreed by 4 people before a decision is considered final.
- Random checks of Crowd’s decisions, and more formal evaluations (in which their decisions are compared with those of expert screeners) show very high accuracy (results consistently demonstrate 99% sensitivity in the identification of RCTs when screening on titles and abstracts).

What do we want the relationship between the crowd, the machine, and reviewers to be?

- Are we expecting the machines and crowds to replace our effort, or some of our effort, as reviewers? Perhaps it’s a restructure of the process. Maybe the machine helps to wade out the noise and then you use the crowd or reviewers to answer the really hard questions or uncertain ones.
- But, do we risk de-skilling staff (or losing incidental learning about the review topic, or refinement of review inclusion criteria) if we remove certain tasks from the systematic review workflow?

We need an agreed evaluation framework of machine/crowd processes

- We want to learn what works and under what circumstances
- Need to better explore what could or should be the role of machine learning (given its evolving)

Methods for updating meta-analysis

Questions about adjusting meta-analyses

- Whether adjustment of updated meta-analyses for Type I error is needed, and if so, when is clear.
- Further meetings with statisticians planned, but for now it's helpful to explore the implications of adjusting and not adjusting in real examples
- If adjustment is used, should this be specified in the protocol, or can it be post-hoc?
- How does the meta-analysis adjustment issue fit with move towards confidence intervals?
- Decision rules and guidance needed to advise authors

Recognition that stopping rules should be informed by other (non-statistical factors)

- Multiple outcomes in a decision framework or subgroup concerns
- Type 1 error relates to a single outcome but decision making relates to broader outcome perspective.
- Issues around including other reasons for uncertainties (e.g. risk of bias or indirectness), do not differ between standard SRs and LSRs.
- Need to distinguish between differences in statistical results and interpretation of reviews

Who does the adjustment and how?

- For now, author teams will need training and/or specialist statistics support, using programs outside of RevMan.
- For living Network Meta Analysis statisticians will need to be fully involved

Intersections with living guidelines*Distinction between a living recommendation and a living guideline*

- May initially focus on a single living recommendation (underpinned by a single LSR)
- But probably less than ideal to have a single living recommendation within a larger static guideline

Feasibility issues for living guidelines

- Living guidelines are affected by more than just changes in the evidence from an LSR (i.e. diagnostic criteria might change; costs and acceptability may change)
- Would also need many LSRs (covering all relevant benefits and harms) to be able to maintain a whole guideline as living

Invites closer collaboration between systematic review and guideline producers

- Collaborations between systematic review producers and guideline developers are strengthening, with closer ties and online platforms, but living evidence invites closer collaboration
- Will be important to engage the guideline groups early to ensure that the LSRs are meeting their needs

Policy and implementation implications*LSRs invite opportunities for deeper collaboration with our stakeholders/review users*

- Given LSRs are particularly suited to priority topics, where the evidence base is growing, they invite closer and deeper collaboration with consumers and other stakeholders, to

ensure they meet their needs. Activities like prioritisation and co-production are particularly important.

- But if we need to build closer relationships, how do we do this on a global level (although this is not an issue limited to LSRs)?

LSRs present a unique marketing opportunity but what is the message we want to sell?

- We need to be clear about what is different and appealing with LSRs (i.e. rapid integration of new evidence without loss of validity)
- Would be helpful to better understand user perspectives on LSRs to inform the messaging.

How do we build the argument to get the ongoing funding needed for LSRs?

- LSRs by their nature require an ongoing commitment, rather than a time-limited commitment, as per most research work.
- We need to be able to demonstrate the efficiency gains of LSRs (without validity or quality loss) and the applicability to guideline developers internationally (effectively a global efficiency).
- Could also make a potential counter argument (i.e. if there wasn't an LSR on this topic, there are costs involved in duplicating evidence synthesis to inform ongoing decisions, OR, if there is no updated review then decision makers may make poor decisions.

Opportunities for consumer engagement

How might consumer engagement opportunities differ with LSRs versus standard SRs?

- LSRs may be more desirable to consumers to be involved in (greater pace of activity and feeding directly into informing practice)
- The more ongoing nature of LSRs invites a longer-term relationship, involving people in all stages/cycles of the LSR (i.e. more coproduction)

Concerns or caveats with engaging consumers in LSRs

- We need to be careful with what we promise about how consumers will be involved
- Would giving consumers greater involvement in LSRs undermine review quality?
- Do we need to understand the impact of consumer engagement in standard systematic reviews first, so we have a baseline to compare to?

Evaluating the process and impact of LSRs

Important impact measures that could be captured for evaluating LSRs

- Cost and resource use
- Review quality (i.e. adherence to PRISMA, ROBIS, etc)

- Time to integrate new studies compared with conventional review updates
- If we could link SRs to particular practice change; then could cost that change and the cost of the SR; may be able to say there are observable differences in the way living recommendations are being used
- Impact on user needs
- Impact on new primary research
- Impact on editorial processes and teams
- Uptake (or speed of uptake) into point-of-care tools for health professionals
- Reduction in unnecessary (redundant) SRs on the same topic
- Learning what components, or types of LSRs are adding value, and in what contexts
- How readers are interacting with LSRs on Cochrane Library (i.e. which components of the reviews are being viewed, downloaded)

Additional considerations for evaluating LSRs

- Evaluating impact is important for securing funding, but already difficult to do for standard SRs
- Must not forget process measures: we don't have a clear sense of what LSRs are yet (in terms of how they are being operationalised). Part of a process evaluation would also generate a clear definition of what an LSR is (or is not) and how it differs from a traditional systematic review.
- Need to identify who are target users and groups are (i.e. LSR production teams, patients, research scientists, practitioners, policy makers and systems) and identify impact metrics for each
- May need to involve an ethicist (there could be ethical dimensions to LSRs that should be explored)

Towards a research agenda

The group proposed the following questions that could form the basis of a research agenda in LSRs:

- What is the value proposition of LSRs for stakeholders/users?
- How should the findings of LSRs be communicated (and what components of the review are most important)?
- How, when and why LSRs should involve consumers and other stakeholders?
- How can LSRs avoid waste and increase efficiency in the evidence ecosystem?
- What are implications of LSRs for policy makers, health systems and guideline developers?

- What methods (and evaluation framework) is needed to fully evaluate pilot LSRs?
- Which steps in the LSR process apply to which topics or situations?
- What should the decision rules be for (1) when to include new evidence (and which evidence) and (2) if, when and how to adjust the meta-analysis for Type I error?

Appendix: Organising committee, speakers, facilitators and attendees

Organising committee

Elie Akl	Holger Schünemann
Julian Elliott	Anneliese Synnot
Jeremy Grimshaw	Tari Turner
Nancy Santesso	

Invited speakers

Elie Akl	Harriet MacLehose	Katrina Sullivan
Jeremy Grimshaw	Anna Noel-Storr	Anneliese Synnot
Julian Elliott	Philippe Ravaud	Laura Weeks
Sarah Elliott	Mark Simmonds	

Session chairs and facilitators

Anna Noel-Storr	Jeremy Grimshaw	Mark Simmonds
Anneliese Synnot	Julian Elliott	Nancy Santesso
Elie Akl	Katrina Sullivan	Philippe Ravaud
Harriet MacLehose	Laura Weeks	Robin Featherstone
Holger Schünemann	Mark Helfand	Sarah Elliott
Ian Shemilt		

Attendees

Alanna Danilkewich	Homa Keshavarz	Meghan Sebastianski
Alexis Turgeon	Joerg Meerpohl	Nigar Sekercioglu
Anna Sawka	Jordi Pardo Pardo	Pasqualina Santaguida
Avtar Lal	Kate Mercer	Peter Tugwell
Ayed Asiri	Kristine Newman	Reed Siemieniuk
Caroline Perron	Laura Banfield	Samantha Craigie
Catherine Boden	Laurie Morrison	Sarah Muir
Cathy Egan	Leslea Peirson	Sathya Karunanathan
Diana Sherifali	Lina Santaguida	Suelen M. Goes
Donna Fitzpatrick	Manisha Pahwa	Yaser Meeralam
Francisco Grajales	Marie-Joëlle Cossi	