Reviewing Elements to inForm RE-running SearcHes: REFRESHing systematic literature reviews
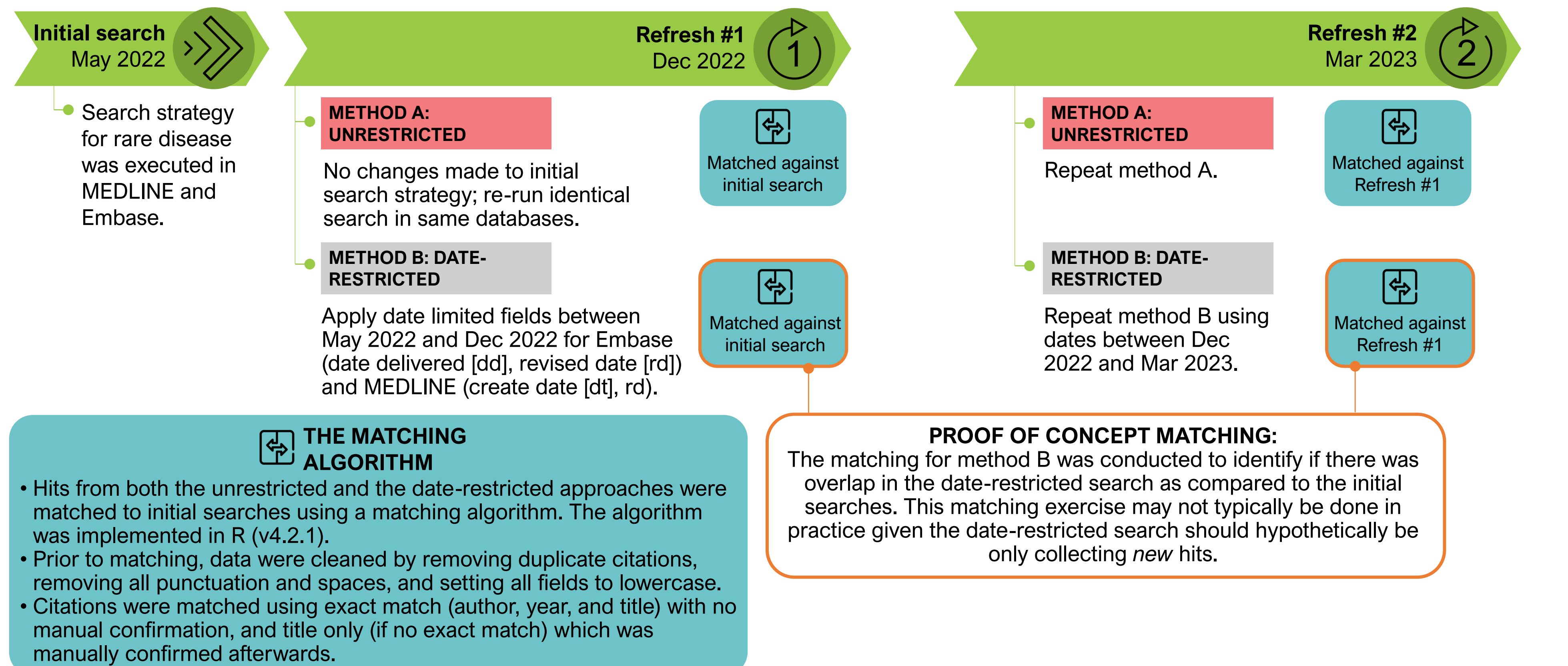
Besada M[1], Deighton A[1], Popoff E[1], Miller SD[1], Friesen M[1], Mickle AT[1], Lozano-Ortega G[1]

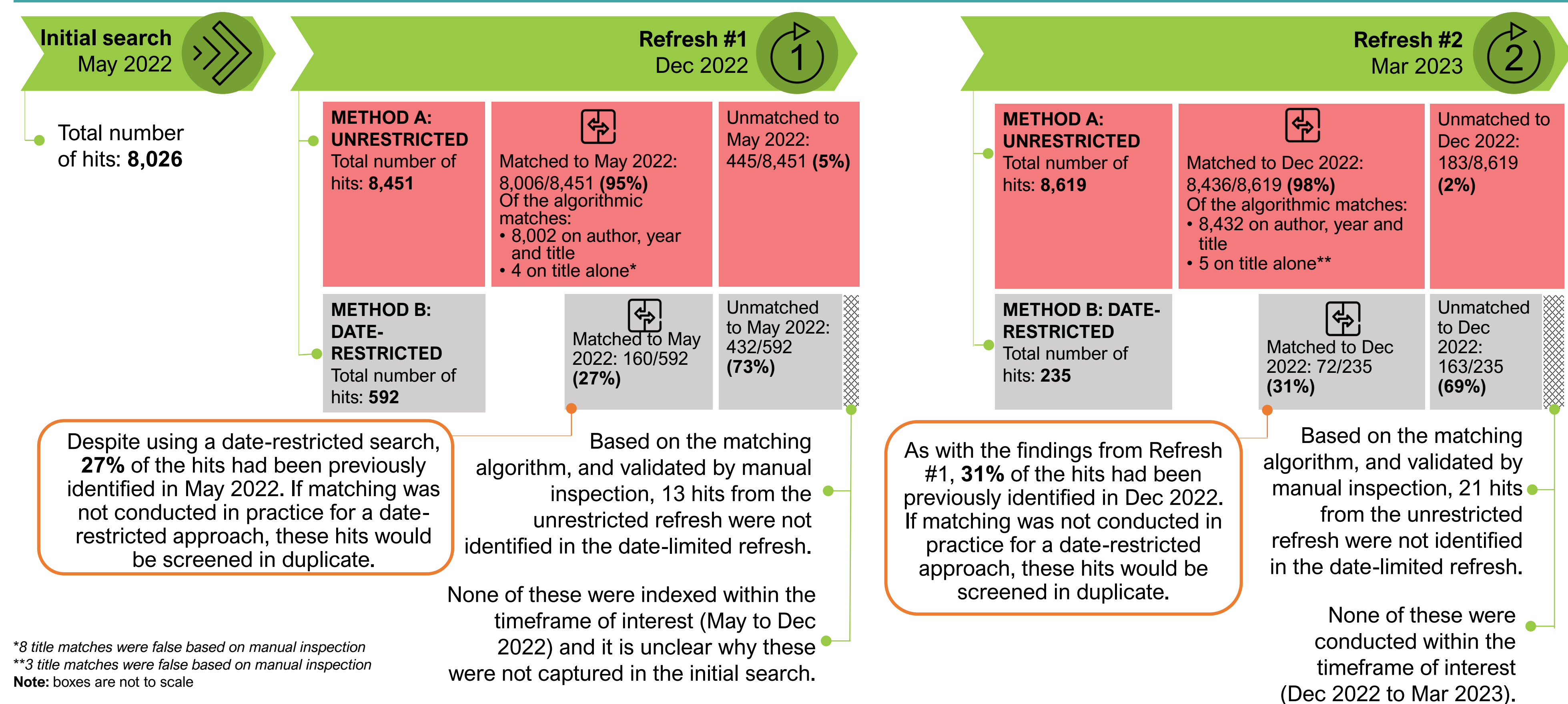[1]Broadstreet HEOR, 201 – 343 Railway St, Vancouver BC Canada

BACKGROUND

- Systematic literature reviews (SLRs) aim to capture all published evidence that meet the study-specific selection criteria (i.e., the Population, Intervention, Comparator, Outcomes and Study Design [PICOS] criteria).[1]
- The majority of SLR evidence is identified through applying search criteria to key databases (e.g., MEDLINE, Embase).
- Re-running searches to stay up to date with the current state of evidence is becoming increasingly important; however, robust methods must be applied.
- To avoid unnecessary duplication in abstract screening, date restriction may be applied to search refreshes; it is unclear if this approach captures all relevant abstracts.
- Alternatively, the search can be re-run and a matching algorithm can be applied to detect and remove any abstracts identified in the original search, eliminating the need to apply date restriction.
- A consensus on the most efficient method that comprehensively identifies new evidence while leaving out previously captured information has yet to be established.

OBJECTIVE

- To compare two approaches for refreshing database searches in OVID using a case study.
- The two approaches were:
  1. A date-restricted approach (with optional matching)
  2. An unrestricted approach with mandatory matching

METHODS

Initial search May 2022
Search strategy for rare disease was executed in MEDLINE and Embase.

Refresh #1 Dec 2022

METHOD A: UNRESTRICTED
No changes made to initial search strategy; re-run identical search in same databases.

Matched against initial search

METHOD B: DATE-RESTRICTED
Apply date limited fields between May 2022 and Dec 2022 for Embase (date delivered [dd], revised date [rd]) and MEDLINE (create date [dt], rd).

Matched against initial search

Refresh #2 Mar 2023

METHOD A: UNRESTRICTED
Repeat method A.

Matched against Refresh #1

METHOD B: DATE-RESTRICTED
Repeat method B using dates between Dec 2022 and Mar 2023.

Matched against Refresh #1

THE MATCHING ALGORITHM
- Hits from both the unrestricted and the date-restricted approaches were matched to initial searches using a matching algorithm. The algorithm was implemented in R (v4.2.1).
- Prior to matching, data were cleaned by removing duplicate citations, removing all punctuation and spaces, and setting all fields to lowercase.
- Citations were matched using exact match (author, year, and title) with no manual confirmation, and title only (if no exact match) which was manually confirmed afterwards.

PROOF OF CONCEPT MATCHING:
The matching for method B was conducted to identify if there was overlap in the date-restricted search as compared to the initial searches. This matching exercise may not typically be done in practice given the date-restricted search should hypothetically be only collecting *new* hits.

RESULTS

Initial search May 2022
Total number of hits: 8,026

Refresh #1 Dec 2022

METHOD A: UNRESTRICTED
Total number of hits: **8,451**

Matched to May 2022: 8,006/8,451 **(95%)**
Of the algorithmic matches:
- 8,002 on author, year and title
- 4 on title alone*

Unmatched to May 2022: 445/8,451 **(5%)**

METHOD B: DATE-RESTRICTED
Total number of hits: **592**

Matched to May 2022: 160/592 **(27%)**

Unmatched to May 2022: 432/592 **(73%)**

Refresh #2 Mar 2023

METHOD A: UNRESTRICTED
Total number of hits: **8,619**

Matched to Dec 2022: 8,436/8,619 **(98%)**
Of the algorithmic matches:
- 8,432 on author, year and title
- 5 on title alone**

Unmatched to Dec 2022: 183/8,619 **(2%)**

METHOD B: DATE-RESTRICTED
Total number of hits: **235**

Matched to Dec 2022: 72/235 **(31%)**

Unmatched to Dec 2022: 163/235 **(69%)**

Despite using a date-restricted search, **27%** of the hits had been previously identified in May 2022. If matching was not conducted in practice for a date-restricted approach, these hits would be screened in duplicate.

Based on the matching algorithm, and validated by manual inspection, 13 hits from the unrestricted refresh were not identified in the date-limited refresh. None of these were indexed within the timeframe of interest (May to Dec 2022) and it is unclear why these were not captured in the initial search.

As with the findings from Refresh #1, **31%** of the hits had been previously identified in Dec 2022. If matching was not conducted in practice for a date-restricted approach, these hits would be screened in duplicate.

Based on the matching algorithm, and validated by manual inspection, 21 hits from the unrestricted refresh were not identified in the date-limited refresh. None of these were conducted within the timeframe of interest (Dec 2022 to Mar 2023).

*8 title matches were false based on manual inspection
**3 title matches were false based on manual inspection
Note: boxes are not to scale

CONCLUSION

An unrestricted approach employing an automated high quality matching algorithm is recommended when refreshing systematic literature reviews for the following reasons:
- An unrestricted approach is more sensitive than a date-restricted approach.
  - In this case study, 13 hits in refresh #1 and 21 hits in refresh #2 would have not been captured had a date-restricted approach been used.
  - A more sensitive approach is particularly important in the context of a scarce evidence base (e.g., a rare disease) where an unrestricted approach may help ensure all available data are captured, even hits outside of the date range not previously identified.
- If the scope of the SLR is expanded (e.g., a new treatment comparator is added), search terms may need to be added to the search strategy.
  - An unrestricted approach will capture all evidence for the newly added terms (from outside of the date limits) as well as the new data from the pre-existing treatments.
  - The unrestricted approach would only require one search, while a date-restricted search would require two separate searches: one unrestricted approach for the newly added terms and a date-restricted approach for the existing terms.

A date-restricted approach may be more appropriate if there are differences in the study record formatting across searches (e.g., different de-duplication strategies, different Endnote and Excel import and export filters). In these situations, the matching algorithm is less likely to capture all matches due to discrepancies in the way authors, years and titles are listed. **If a date-limited approach is preferred, the following considerations should be made:**
- Multiple date fields should be included to ensure comprehensive results.
  - However, it should be noted that the create date (dt) is the date when the records were added to PubMed, not the date of publication.
  - The publication date fields in Ovid databases consist of non-standard values supplied by publishers (e.g., 15-January-2010, Jan-Feb 2014, Winter 2016) and are unsuitable for limiting by a date range more precise than publication year.[2]
- Though matching is optional in a date-restricted approach, it prevents duplicating work.
  - Although matching may be considered an additional step in the SLR process, in this case study, a non-negligible number of hits in the date-restricted search had been previously identified.
  - Without matching, the date-restricted approach may identify previously screened hits, resulting in duplicated work.

ADDITIONAL CONSIDERATIONS

Living systematic reviews
A particular case of SLR refresh is the living systematic review (LSR) which aims to continuously monitor emerging evidence using regular refreshes at pre-specified, short intervals.[1]

LSRs are subject to further considerations: in addition to the two methods considered here, reviewers can opt for employing auto-alerts as outlined in the Cochrane LSR guidelines.[1] Auto-alerts may be more efficient as the searches and outputs are generated automatically at pre-specified intervals. However, the output of auto-alerts (a text file with a list of abstracts) may be unsuitable for evidence bases that are large and require a more robust system for citation management.

Future work
In addition to database searches, future work for establishing robust methods for other steps in the refresh process comprise:
- Different considerations for a LSR versus discrete SLR refresh.
- Presenting data in one versus multiple Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagrams.[3]
- Defining go/no go decision for publishing new results.

REFERENCES

1. Cochrane. Guidance for the production and publication of Cochrane living systematic reviews: Cochrane Reviews in living mode. 2019.
2. Wolters Kluwer. How Can I Limit Search Results By A Date Range In Ovid? ND.
3. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71