Journal of Clinical Epidemiology 173 (2024) 111466

Journal of Clinical Epidemiology

ORIGINAL RESEARCH

Guidance on how to efficiently find, choose, and use available systematic reviews was developed

Tim Mathes[a,*], Peggy Prien[b], Irma Klerings[c], Hannah Ewald[d], Corinna Dressler[e], Thomas Harder[f], Fülöp Scheibler[g,h], Roland Büchter[i], Cordula Braun[j], Kathrin Grummich[k], Michaela Eikermann[l], Corinna Schaefer[b], Dawid Pieper[m,n,o]

[a]*Department of Medical Statistics, University Medical Center Goettingen, Humboldtallee 32, D-37073 Göttingen, Germany*
[b]*German Agency for Quality in Medicine, Berlin, Germany*
[c]*Cochrane Austria, Department for Evidence-based Medicine and Evaluation, Danube University Krems, Krems, Austria*
[d]*University Medical Library, University of Basel, Basel, Switzerland*
[e]*Charité – Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Medical Library, Berlin, Germany*
[f]*Robert Koch Institute, Immunization Unit, Berlin, Germany*
[g]*National Competency Center for Shared Decision Making, University Hospital Schleswig-Holstein, Kiel, Germany*
[h]*SHARE TO CARE. Patientenzentrierte Versorgung GmbH, Cologne, Germany*
[i]*Institute for Research in Operative Medicine (IFOM), Faculty of Health, School of Medicine, Witten/Herdecke University, Cologne, Germany*
[j]*Hochschule 21, Buxtehude, Germany*
[k]*ZB MED - Information Centre for Life Sciences, Cologne, Germany*
[l]*Federal Medical Advisory Service, Department of Evidence Based Medicine, Essen, Germany*
[m]*Faculty of Health Sciences Brandenburg, Brandenburg Medical School (Theodor Fontane), Institute for Health Services and Health System Research, Rüdersdorf, Germany*
[n]*Center for Health Services Research, Brandenburg Medical School (Theodor Fontane), Rüdersdorf, Germany*
[o]*Evidence Based Practice in Brandenburg: A JBI Affiliated Group, University of Adelaide*

Accepted 9 July 2024; Published online 15 July 2024

Abstract

Objectives: The aim of this paper is to provide clinicians and authors of clinical guidelines or patient information with practical guidance on searching and choosing systematic reviews(s) (SR[s]) and, where adequate, on making use of SR(s).

Study Design and Setting: At the German conference of the Evidence-Based Medicine Network (EbM Network) a workshop on the topic was held to identify the most important areas where guidance for practice appears necessary. After the workshop, we established working groups. These included SR users with different backgrounds (eg, information specialists, epidemiologists) and working areas. Each working group developed and consented a draft guidance based on their expert knowledge and experiences. The results were presented to the entire group and finalized in an iterative process.

Results: We developed a practical guidance that answers questions that usually arise when choosing and using SR(s). (1) How to efficiently find high-quality SRs? (2) How to choose the most appropriate SR? (3) What to do if no SR of sufficient quality could be identified? In addition, we developed an algorithm that links these steps and accounts for their interaction. The resulting guidance is primarily directed at clinicians and developers of clinical practice guidelines or patient information resources.

Conclusion: We suggest practical guidance for making the best use of SRs when answering a specific research question. The guidance may contribute to the efficient use of existing SRs. Potential benefits when using existing SRs should be always weighted against potential limitations.

Keywords: Systematic reviews; Redundant reviews; Literature searching; Study selection; Evidence synthesizes; Evidence-Based-Medicine

1. Introduction

Clinical decisions that are based on the results of individual primary studies may be misleading. Therefore,

* Corresponding author. Department of Medical Statistics, University Medical Center Göttingen, Humboldtallee 32, D-37073 Göttingen, Germany.
E-mail address: tim.mathes@med.uni-goettingen.de (T. Mathes).

https://doi.org/10.1016/j.jclinepi.2024.111466

What is new?

Key findings

- We provide practical guidance, on searching and choosing systematic review (SR[s]) and, where adequate, on making use of SR(s), including an algorithm for integrating these steps. The guidance is not a one size fits all approach but allows some flexibility to acknowledge that resource limitations and required timelines can vary between evidence synthesis types and settings.

What this adds to what is known?

- To our knowledge, this is the first such guidance.

What is the implication and what should change now?

- The suggested guidance could facilitate and harmonize the choice of SR(s).

systematic reviews (SRs) are key to informing health-care decisions. Especially when many research questions need to be answered and time and resources are limited, the use of available SRs can be an efficient approach. Using only the most appropriate SR instead of all that meet the inclusion criteria can further reduce the time and resources needed. Moreover, it can avoid challenges (eg, handling overlapping SRs), and confusion (eg, conflicting results) that may arise when including multiple SRs.

Several manuals provide guidance on searching for, selecting, and synthesizing SRs (eg, Pollock 2022 [1], Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften (AWMF) 2020 [2]). However, these manuals do not address how to make the best use of available SRs. Lunny et al recently performed a scoping review on tools for choosing one SR from multiple conflicting or discordant SRs on the same research question [3,4]. They only identified one tool developed for this purpose—the Jadad algorithm. In a replication study, this tool proved to be poorly operationalized and unreliable [4]. Therefore, the authors recommended that ''health-care providers, policymakers, patients, and researchers address conflicts between review findings by choosing the SR(s) with meta-analysis of ranomdimzed controlled trials that most closely resemble(s) their clinical, public health, or policy question, is/are the most recent, comprehensive (ie, the number of included ranomdimzed controlled trials), and at the lowest risk of bias.'' Lunny et al did not operationalize these criteria further and only considered the situation of conflicting and discordant SRs, ie, the general situation of having more than one SR on the same research question was not addressed. Pilic et al published guidance on how to choose SRs when developing evidence-based vaccination recommendations. The guidance covers assessing the relevance to the research question, up-to-dateness and methodological quality, as well as the transferability and applicability of the findings to other populations or settings. However, no suggestions are made for integrating the different aspects [5]. In addition, often situations arise, where users end up with an array of outdated or low to medium quality SRs, which raises the question, whether, under which circumstances and how these can be used to inform decisions (in the absence of high-quality, up-to-date SRs). Notably, there is some indication that, at least in certain areas of research, a considerable number of SRs are of low methodological quality. Eg, in the Systematic Reviews on Vaccines (SYSVAC) database, a repository of SRs on vaccination topics hosted by World Health Organization, more than 90% of reviews on efficacy/effectiveness or safety are rated as being of critically low quality according to AMSTAR-2 (A. Pilic, personal communication).

The aim of our project was to provide practical guidance on searching and choosing SR(s) and, where adequate, on making use of SR(s) that suffer minor to moderate methodological limitations.

The resulting guidance primarily aims to support clinicians and developers of clinical practice guidelines or patient information resources who understand the core principles of evidence-based medicine (EbM) and SR methodology. Thus, it may also be used by other groups who have the necessary knowledge including practitioners, decision-makers, researchers justifying or informing a new study, and EbM teachers.

In everyday practice, time and resources are usually limited. Thus, our objective was to describe how to best use existing SRs given the available time and resources, while being as comprehensive as possible. By ''best use'' we mean identifying the most up-to-date, the best (ie, with the highest methodological quality), and largest (ie, with the most included relevant studies) SR across all available SRs. We do not suggest a strict one-size-fits-all-approach because decisions on how to identify and choose SRs and to make use of their results will vary depending on individual resources, settings, and health-care questions.

2. Approach for developing the guidance

This was a multiple-step project. At the Conference of the Network for EbM in Lübeck, Germany in 2022 a workshop on the topic was held to discuss the issue and identify the most important areas where guidance for practice appears necessary [6]. The participants (about 30) were mainly people who are involved in preparing evidence syntheses for informing decisions such as preparing guidelines and patient information resources from various professional disciplines, including health professionals, methodologists, medical journalists, sociologists, and public health researchers.

Three main questions emerged from the workshop include the following:

1. How to efficiently find high-quality SRs?
2. How to choose the most appropriate SR?
3. What to do if no SR of sufficient quality could be identified?

After the workshop, we established working groups for each of these questions. The 3 working groups included 13 experts in total, and each group included at least five persons. Note, that some experts contributed to more than one group. The working groups included SR users from Germany, Austria, and Switzerland with different backgrounds (eg, information specialists, epidemiologists, health-care professionals) and from different working areas (eg, clinical practice guideline development, patient information development).

For informing the development of the guidance, all group members contributed the relevant literature on the topic known to them. We did not conduct systematic literature searches but tried to identify all relevant key publications by forward and backward citation searches. In an iterative, but not formalized process, each working group developed and consented a draft guidance based on their expert knowledge and experiences. The results were presented to all members and refined in another iterative process until a consensus in the entire group was reached. Finally, to receive feedback from a wider range of targeted users, the results were presented and discussed during a second workshop at the conference of the EbM Network in Berlin, Germany, in 2024. The participants of this workshop were mostly people (about 15) who had not participated in the previous workshop. During this workshop, the draft guidance was presented and discussed, but no need for further relevant changes was identified.

3. Results

3.1. How to efficiently find SRs?

Conducting systematic database searches for SRs can be time-consuming, can yield thousands of hits, and can be costly when many potentially eligible SRs are not freely available. Furthermore, the identified SRs may be outdated (ie, missing relevant new studies), redundant (ie, addressing the same research question) or of low methodological quality.

A resource-conscious search strategy creates a balance between precision and sensitivity, that, when searching for SRs to answer a narrow health-care question, tends toward precision. Therefore, we suggest a stepwise search approach with a focus on precision. As designing systematic searches requires specific knowledge and skills, it is advisable that a person with expertise and experience in designing search strategies (eg, information specialist) is involved.

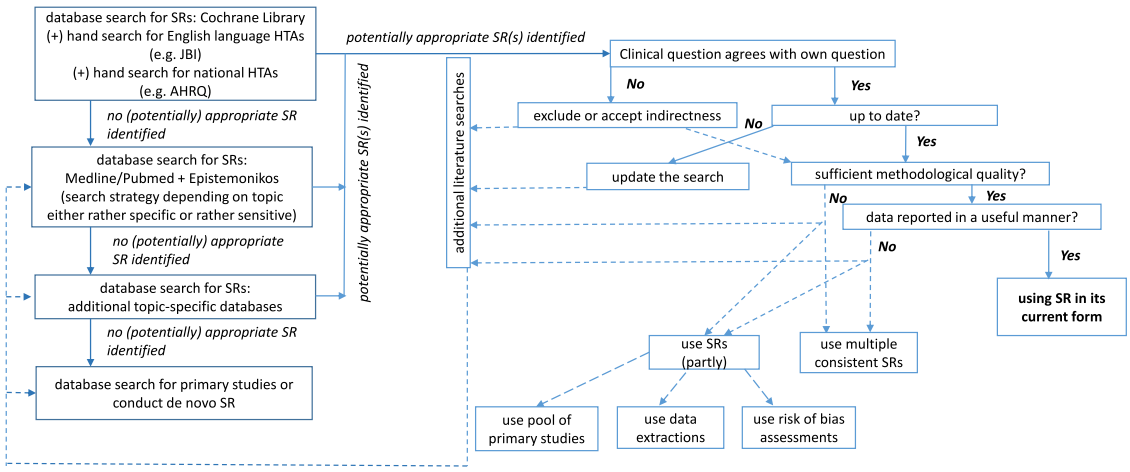
3.1.1. Stepwise search approach

- Step 1: Focused Systematic Search (FSS): as a first step, we suggest searching selected providers known for producing high-quality SRs [7–9]. This could involve searching only the Cochrane Database of Systematic Reviews, possibly combined with searches in English-language (eg, National Institute for Health and Care Excellence [NICE], Agency for Healthcare Research and Quality [AHRQ], Joanna Briggs Institute [JBI]) and non-English Health Technology Assessment (HTA) databases suitable for the respective context (eg, Austrian Institute for Health Technology Assessment [AIHTA], Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen [IQWiG]). SRs from these organizations typically provide a substantial number of subgroup analyses and detailed reporting, increasing the likelihood of adequately answering the specific health-care question. However, critically appraising identified SRs using validated tools, and evaluating their up-to-dateness (see section 3.2.), is crucial. If one of the identified SRs meets all key criteria for selection (see Fig), a comprehensive systematic search may no longer be required. Our definition of an FSS explicitly excludes the use of SR(s) selectively, which are provided by individual stakeholders (eg, guideline panel members). To assure reproducibility, each FSS requires a detailed documentation (including websites, click paths, filters, search terms, etc.).
- Step 2: Comprehensive Systematic Search in databases: if the FSS does not yield any relevant results, a more comprehensive search is required [10]. This involves searching several (at least 2) databases, with adapted search strategies and filters for SRs. Eg, searching Medline in combination with Epistemonikos has shown to achieve a very high sensitivity for identifying SRs, provided that search strategies are adequately applied [11,12]. Furthermore, Epistemonikos includes references identified in subscription-based databases such as Embase and CINAHL. For highly specific research questions, or when no adequate results have been retrieved, searches of subject-specific databases (eg, PEDro) may be considered (for database lists see Search Smart [13]).

3.1.2. Tips for search strategy development

If the research question is broad, if initial exploratory searches yield too many results, or if a high number of results is expected, we suggest a precision-focused search strategy and considering the following options:

- For SRs of interventions, a precision-focused search usually combines at least the PICO (patients/population, intervention, comparison, outcome) components P and I. For other types of research questions (eg, diagnosis, prognosis), this is usually analogous (eg, combining P with index test or prognostic factor).
- Searching only in title fields yields significantly fewer results [14]. This approach is particularly suitable for



+ additionally; (+) optionally; solid lines: fixed next step; dashed lines: optional next step requiring balancing between accepting limitations and available time and resources
SR systematic review; HTA: Health Technology Assessment; AHRQ Agency for Healthcare Research and Quality; JBI Joanna Briggs Institute;

Figure. Algorithm illustrating the overall process on how to efficiently find, choose, and use available systematic reviews.

search terms that are a secondary aspect of other topics and therefore often appear in abstracts (eg, depression as a secondary outcome). This strategy can be applied to specific keywords or to all keywords of the search. We suggest using as many synonyms as possible by checking, eg, reference lists, abstracts of other relevant SRs, involving experts as well as using truncation to search for word variations.

- Database interfaces/platforms may provide options to increase search precision, eg, limiting subject headings to "major focus" or the use of frequency operators to define a minimum number of times a term must appear in an abstract [14]. Similarly, proximity operators can be helpful for further increasing the precision of the search.
- Choose a short search period. Older SR(s) may be outdated. Especially if recent primary studies are known, the search period can be narrowed accordingly.
- The more time is invested in developing a precise search strategy, the more time can potentially be saved during the screening of titles and abstracts [15]. However, a saturation point may be reached when developing the search strategy takes longer than the screening would take.

3.1.3. Citation-based searching

There is a risk of drawing wrong conclusions if pertinent SRs are missed that may provide contradictory results. Normally, high-quality SRs that have applied adequate methods should lead to similar findings [16]. However, this may not always be the case due to diverging inclusion criteria, search methods or decisions about whether or not to synthesize results. To reduce the risk of missing high-quality SRs with a precise search strategy, the following measures may be taken into account:

- Co-citing: If a potentially appropriate SR is identified, check for other SRs that include at least one of the primary studies cited [17].
- Forward citation searching: If you know recent relevant primary studies (“landmark trials”), use these for forward citation searching. Check for recent SRs among the results.
- Involving clinical experts: Experts are usually familiar with the literature in their field. If they find that relevant evidence is missing, adjust the search strategy.

3.2. How to choose the most appropriate SR(s)?

If potentially eligible SR(s) are available, their appropriateness needs to be assessed. Ideally, one or more appropriate SR(s) are identified that can be used.

When selecting appropriate SR(s) and choosing the most appropriate SR(s), several criteria can be relevant.

3.2.1. Criteria for choosing an SR

3.2.1.1. Research question. SRs that, at first sight, appear to address the same research question often differ in the details of their research question or selection criteria. When selecting appropriate SRs it is important to check that the eligibility criteria of the identified SRs match one’s research question because SRs that do not exactly meet the selection criteria of the research question may provide no relevant answer. More specifically, narrow eligibility criteria resulting in the exclusion of primary studies relevant to the question of interest may lead to exclusion of the SR. This can occur when an SR is limited to subgroups,

specific outcomes or specific settings. In contrast, when the research question of interest is reported in a separate analysis, it can be used conveniently. If identified SRs have broader eligibility criteria and do not separately report the research question of interest, assessing the indirectness and transferability to the health-care question can help to decide whether and how the SR can be used [18].

3.2.1.2. Up-to-dateness. Several fixed thresholds for considering a literature search up-to-date have been suggested. Eg, AMSTAR 2 considers 24 months to be appropriate [19]. However, we believe that assessing the up-to-dateness should not simply be based on an arbitrary cut-off date [20]. Rather, the likelihood that new studies will change the findings or the quality of evidence should be considered [21]. If high-quality evidence is available or if no new study results are expected to be published, even older searches may be considered up-to-date. To assess whether new studies have been published since the last search of the SR(s), abbreviated literature searches may be helpful.

3.2.1.3. Methodological quality. AMSTAR 2 or ROBIS (with AMSTAR 2 developed for SRs of interventions) are validated tools for assessing the methodological quality and risk of bias of SRs and may be used to select adequate SRs [19,22]. However, not all domains of these tools are of equal importance when choosing an existing SR for a specific decision-making situation. An adequate literature search is often a decisive factor because it determines the completeness of the pool of primary studies and is therefore a premise for reusing the data of the included SR(s) (see section 3.3). In addition, performing and presenting a risk of bias assessment is often an indicator for a good quality SR [23]. However, other items of critical appraisal tools like AMSTAR 2 can be of equal importance in specific contexts or situations. If, eg, a guideline group uses a structured approach for synthesizing the data (eg, GRADE [24]) from an identified SR, adequate reporting of the results will be required. In contrast, it will not be important that the discussion section of an identified SR is appropriate.

3.2.1.4. Reporting of results/data in a useful manner. SR(s) are only useful if they report the results or data in a useful manner. What this means can vary depending on purpose. Eg, if a guideline group decides to provide summaries of findings tables for each research question, an SR needs to contain the tables or the necessary data with which a summary of findings table can be prepared. Nonadherence to reporting guidelines such as PRISMA [25] or its derivates is usually a red flag for inadequate reporting. Limiting the reporting of effect estimates to relative effects or apparent inconsistencies in the data can be an indicator for potentially inadequate data [26].

3.2.2. Approach for applying the criteria (research question, up-to-dateness, methodological quality, reporting of results)

We suggest a 2-step approach for determining the included SR(s). The first step is to identify potentially appropriate SRs efficiently by applying a set of knockout criteria, ie, specifying minimal requirements for inclusion using the criteria above. The second step is to apply the same criteria in more depth to all remaining SR(s) simultaneously for choosing one, or, if one SR does not fully cover the research question or has other limitations, a few SR(s).

We suggest applying the criteria in both steps in the following order (see also Fig) because some criteria are clearly not useful (eg, data are reported adequately) when other criteria (eg, up-to-dateness) are not met.

1. Assessing agreement with the research question (directness)
2. Assessing up-to-dateness
3. Assessing methodological quality
4. Assessing adequacy of reported results

In addition, some criteria can be applied more easily. Therefore, checking against these criteria first will be more efficient. Eg, broad agreement with the own research question can usually be checked based on information from the abstract. If none of the identified SRs satisfies all criteria, further strategies related to how to deal with these SRs may be considered (see section 3.3.). How strictly the criteria should be formulated depends on the expected number of potentially appropriate SRs and the available time and resources. The more potentially appropriate SRs are available, and the more resources are available for additional tasks (eg, additional searches, performing new data extractions) the stricter the knockout criteria can be (and vice versa).

If more than one SR pass the minimal knockout criteria, a more thorough application of the criteria is needed to identify which potentially appropriate SRs are actually appropriate. Usually, the SRs will satisfy the different criteria to varying degrees. Eg, one SR might be the most up-to-date but may at the same time be of lower methodological quality than another. In addition, when deciding which SR to choose, someone will usually weight the criteria differently depending on the research question and decision context. Eg, up-to-dateness may be considered more important in a dynamic research field. Thus, when applying the knockout criteria in depth, we suggest doing so across all potentially appropriate SRs

simultaneously and keeping in mind that the strength of one SR may be the weakness of another SR and vice versa. This balancing of the criteria implies simultaneous value judgments and hence some subjectivity. At this stage, either a single or several SRs may be chosen. Choosing more than one SR can be useful, when, eg, a single SRs does not fully address the own research question (eg, not all relevant outcomes) or has limitations concerning the other criteria. In addition, it can be useful for assessing the robustness of the results and explore possible reasons (eg, eligibility criteria, methodological quality) for differences (see next section).

3.3. What to do if no SR of sufficient quality can be identified?

Sometimes, only SRs that are of insufficient methodological quality, or outdated, or suffer from problems with reporting, might be identified. Depending on resources and the research question, it may be helpful to check whether some elements (eg, the identified pool of primary studies, data extractions, or risk of bias assessments) of the identified low quality SRs can still be used to answer the health care question before conducting systematic searches for primary studies.

3.3.1. Using the identified pool of primary studies

This can be an option if one is very confident that the identified pool of primary studies in a review is based on a sufficiently comprehensive search. Deciding whether a literature search was sufficiently comprehensive can be challenging. We suggest that authors use the PRESS guideline to assess about the comprehensiveness of the literature search [27]. Involving a person with specialist expertise and experience in literature searching is advisable at this step. This will help to decide whether the search will likely satisfy the required sensitivity and specificity. In addition, authors need to check whether the identified pool of primary studies is comprehensive or whether there are relevant studies missing. For this task, a sufficient level of content expertise is required. The absence of certain studies can have several reasons, such as date of search or eligibility criteria.

If only multiple low or moderate quality SRs are available, authors may also consider comparing the pool of studies across reviews. The authors are likely to face differences between the pool of primary studies, eg, due to differences in the eligibility criteria or search dates. It may be helpful to check the reason why a primary study was excluded by inspecting the list of excluded studies, if available. Checking this for multiple available SRs will take time. However, it could increase the confidence in a study pool if no major differences can be detected. It is also important to note that differences in the pool of studies are more likely to be observed in emerging fields, such as COVID-19, as the date of the search as well as the information sources (eg, preprint servers) searched can have an impact.

If it is likely that new evidence has become available that was not included in the identified SRs, it may be helpful to consider using the pool of primary studies and updating it. This will usually involve searching for new studies that are available after the most recent search was conducted. If authors decide to not update the search, a clear rationale should be provided.

A prerequisite for both steps (ie, comparing pools of studies and updating the search) is that the literature search is adequately reported. To reproduce and update a search, all details need to be known. It is likely that this will not be the case for most SRs as reporting is frequently insufficient [28,29].

3.3.2. Using data extractions

Data extraction errors in SRs are common [30]. In general, one should be cautious when using extracted data from other SRs. At least some quality checks should be performed. If there are multiple low or moderate quality SRs available, the extracted data should be compared and checked for inconsistency across these SRs. If only one SR is available, one should extract data from a sample of the included studies and check these against the data presented in this SR. The latter step may also be performed in addition to checking consistency across multiple SRs. These plausibility checks will allow to judge whether data extraction was accurate.

Since some types of information may be more relevant than others, it can be pragmatic to distinguish between the types of extracted data. Eg, an error in the information on study setting may have less impact on formulating recommendations than relying on incorrect effect estimates.

3.3.3. Using risk of bias assessments

Risk of bias assessments can differ for the same studies across SRs [31−33]. Thus, we advise against the uncritical use of risk of bias assessments made by SR authors. In general, the same approach as for data extraction (see above) can be adopted. Obviously, authors need to make sure that the instrument used in an SR is the same as they intend to use in their own SR.

3.3.4. Using multiple insufficient SRs

In the case of multiple available SRs, authors may consider checking the conclusions of these reviews. If all SRs come to the same conclusion, irrespective of differences in the pool of primary studies, extracted data and risk of bias assessments, it may increase the confidence in the results. Nevertheless, a minimum set of criteria (see knockout criteria section 3.2.) for the methodological quality of the SRs should be defined beforehand.

4. Conclusion

Based on our experience, we suggest a layered approach that accounts for several possible constraints including time/resources for making the best use of SRs when

seeking to answer a specific research question (eg, in a clinical practice guideline).

The approach mainly considers the agreement of the own research question and the SR's research question, up-to-dateness, methodological quality, and reporting the necessary data when choosing an SR(s). Such practical guidance on the efficient use of existing SRs appears particularly warranted in the light of the rapid growth of the number of SRs published each day and the consequential additional effort when including all SRs on the same topic [34]. Some steps of the suggested approach may be further facilitated in the future by using artificial intelligence (eg, literature search, applying knockout criteria for study selection). However, it is important to note that these recommendations are based on expert discussions. A more formalized approach (eg, a systematic literature review followed by a Delphi process) may lead to a more nuanced guidance.

Despite this limitation, the benefit of following our suggested strategies can be that a de novo SR might be avoided and thus resources can be used more efficiently. However, depending on the context and search results, using existing SRs may be inferior compared to conducting a new SR because some limitations must be considered: The research question may not always be fully suitable, data may not be provided appropriately, or additional efforts may be necessary before using the SR(s) (eg, updating searches, extracting data). The additional effort for revising an inappropriate SR can be even larger than conducting a new SR from the scratch. Therefore, potential advantages when using existing SRs should be always weighed against potential disadvantages compared to a new SR beforehand.

CRediT authorship contribution statement

Tim Mathes: Writing – original draft, Visualization, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Peggy Prien:** Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Irma Klerings:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Hannah Ewald:** Writing – review & editing, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Corinna Dressler:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Thomas Harder:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Fülöp Scheibler:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Roland Büchter:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Cordula Braun:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Kathrin Grummich:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Michaela Eikermann:** Writing – review & editing, Methodology, Investigation, Formal analysis, Conceptualization. **Corinna Schaefer:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Dawid Pieper:** Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization.

Data availability

No additional data generated.

Declaration of competing interest

There are no competing interests for any author.

References

[1] Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Chapter V: overviews of reviews. In: Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al, editors. Cochrane Handbook for Systematic Reviews of Interventions version 63. Cochrane; 2022. (updated February 2022). Available at: www.training.cochrane.org/handbook. Accessed July 30, 2024.

[2] Arbeitsgemeinschaft der Wissenschaftlichen Medizinischen Fachgesellschaften - Institut für Medizinisches Wissensmanagement. Ärztliches Zentrum für Qualität in der Medizin, Cochrane Deutschland Stiftung, Institut für Evidenz in der Medizin, Institut für Medizinische Biometrie und Statistik, Freiburg. Manual Systematische Recherche für Evidenzsynthesen und Leitlinien. Germany: AWMF; 2020.

[3] Lunny C, Brennan SE, McDonald S, McKenzie JE. Toward a comprehensive evidence map of overview of systematic review methods: paper 1-purpose, eligibility, search and data extraction. Syst Rev 2017;6(1):231.

[4] Lunny C, Thirugnanasampanthar SS, Kanji S, Ferri N, Pieper D, Whitelaw S, et al. How can clinicians choose between conflicting and discordant systematic reviews? A replication study of the Jadad algorithm. BMC Med Res Methodol 2022;22:276.

[5] Pilic A, Reda S, Jo CL, Burchett H, Bastías M, Campbell P, et al. Use of existing systematic reviews for the development of evidence-based vaccination recommendations: guidance from the SYSVAC expert panel. Vaccine 2023;41(12):1968–78.

[6] Pieper D, Mathes T, Prien P, Schaefer C. Zu viel, zu schlecht, zu unbrauchbar? Systematische Übersichtsarbeiten intelligent nutzen. In: EbM-Netzwerk, editor. 23 Jahrestagung des Deutschen Netzwerks Evidenzbasierte Medizin; 01. Lübeck. Düsseldorf: German Medical Science GMS Publishing House; 2022.

[7] Deshpande S, Misso K, Westwood M, Stirk L, de Kock S, Kleijnen J, et al. PRM77 - not all Cochrane reviews are good quality systematic reviews. Value Health 2016;19(7):A371.

[8] Prien P, Schaefer C. Reducing waste through targeted evidence acquisition (Poster). In: 18th GIN Conference; 25. Online-Kongress; 2021.

[9] Vader I, Schorr S, Schaefer C, Härter M. NVL Asthma: Methodische Qualität der eingeschlossenen systematischen Reviews (Präsentation). In: EbM-Netzwerk, editor. 19 Jahrestagung des Deutschen Netzwerks Evidenzbasierte Medizin e V; 8. Graz. Düsseldorf: German Medical Science GMS Publishing House; 2018. 10.3.2018.

[10] Pollock M, Fernandes RM, Newton AS, Scott SD, Hartling L. A decision tool to help researchers make decisions about including systematic reviews in overviews of reviews of healthcare interventions. Syst Rev 2019;8(1):29.

[11] Goossen K, Hess S, Lunny C, Pieper D. Database combinations to retrieve systematic reviews in overviews of reviews: a methodological study. BMC Med Res Methodol 2020;20:138.
[12] Heinen L, Goossen K, Hirt J, Lunny C, Puljak L, Pieper D. Datenbankkombinationen für die Recherche systematischer Reviews: eine aktualisierte methodische Studie. In: EbM-Netzwerk, editor. 24 Jahrestagung des Netzwerks Evidenzbasierte Medizin e V; 22. Potsdam. Düsseldorf: German Medical Science GMS Publishing House; 2023.
[13] Searchsmart. Available at: https://www.searchsmart.org/?~. Accessed October 3, 2023.
[14] Casey M, Raynor M, Jacob C, Sharp S, McFarlane E. Improving the precision of search strategies for guideline surveillance. Res Synth Methods 2020;11(6):903−12.
[15] Bramer WM, de Jonge GB, Rethlefsen ML, Mast F, Kleijnen J. A systematic approach to searching: an efficient and complete method to develop literature searches. J Med Libr Assoc 2018;106:531−41.
[16] Ewald H, Klerings I, Wagner G, Heise TL, Dobrescu AI, Armijo-Olivo S, et al. Abbreviated and comprehensive literature searches led to identical or very similar effect estimates: a meta-epidemiological study. J Clin Epidemiol 2020;128:1−12.
[17] Hirt J, Nordhausen T, Appenzeller-Herzog C, Ewald H. Citation tracking for systematic literature searching: a scoping review. Res Synth Methods 2023;14(3):563−79.
[18] Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines: 8. Rating the quality of evidence–indirectness. J Clin Epidemiol 2011;64:1303−10.
[19] Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. Amstar 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:j4008.
[20] Gillian S, Katy S, James T. Is a one-size-fits-all '12-month rule' appropriate when it comes to the last search date in systematic reviews? BMJ Evidence-Based Medicine 2023;28(6):359.
[21] Garner P, Hopewell S, Chandler J, MacLehose H, Schünemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ 2016;354:i3507.
[22] Whiting P, Savović J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: a new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol 2016;69:225−34.
[23] Leclercq V, Hiligsmann M, Parisi G, Beaudart C, Tirelli E, Bruyère O. Best-worst scaling identified adequate statistical methods and literature search as the most important items of AMSTAR2 (A measurement tool to assess systematic reviews). J Clin Epidemiol 2020;128:74−82.
[24] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383−94.
[25] PRISMA 2020. 2020. Available at: https://www.prisma-statement.org/prisma-2020. Accessed October 10, 2023.
[26] Novelli M, Baldi Antognini A, Boffetta P, Ioannidis JP, Spatari G, Violante FS. Reporting only relative effect measures was potentially misleading: some good practices for improving the soundness of epidemiological results. J Clin Epidemiol 2021;137:195−9.
[27] McGowan J, Sampson M, Salzwedel DM, Cogo E, Foerster V, Lefebvre C. PRESS peer review of electronic search strategies: 2015 guideline statement. J Clin Epidemiol 2016;75:40−6.
[28] Biocic M, Fidahic M, Puljak L. Reproducibility of search strategies of non-Cochrane systematic reviews published in anaesthesiology journals is suboptimal: primary methodological study. Br J Anaesth 2019;122:e79−81.
[29] Koffel JB, Rethlefsen ML. Reproducibility of search strategies is poor in systematic reviews published in high-impact pediatrics, cardiology and surgery journals: a cross-sectional study. PLoS One 2016;11:e0163309.
[30] Mathes T, Klaßen P, Pieper D. Frequency of data extraction errors and methods to increase data extraction quality: a methodological review. BMC Med Res Methodol 2017;17:152.
[31] Könsgen N, Barcot O, Heß S, Puljak L, Goossen K, Rombey T, et al. Inter-review agreement of risk-of-bias judgments varied in Cochrane reviews. J Clin Epidemiol 2020;120:25−32.
[32] Armijo-Olivo S, Ospina M, da Costa BR, Egger M, Saltaji H, Fuentes J, et al. Poor reliability between Cochrane reviewers and blinded external reviewers when applying the Cochrane risk of bias tool in physical therapy trials. PLoS One 2014;9:e96920.
[33] Jordan VM, Lensen SF, Farquhar CM. There were large discrepancies in risk of bias tool judgments when a randomized controlled trial appeared in more than one systematic review. J Clin Epidemiol 2017;81:72−6.
[34] Hoffmann F, Allers K, Rombey T, Helbach J, Hoffmann A, Mathes T, et al. Nearly 80 systematic reviews were published each day: observational study on trends in epidemiology and reporting over the years 2000-2019. J Clin Epidemiol 2021;138:1−11.