Journal Pre-proof

Resource use during systematic review production varies widely: a scoping review

B. Nussbaumer-Streit , M. Ellen , I. Klerings , R. Sfetcu , N. Riva ,
M. Mahmić-Kaknjo , G. Poulentzas , P. Martinez , E. Baladia ,
L.E. Ziganshina , M.E. Marqués , L. Aguilar , A.P. Kassianos ,
G. Frampton , A.G. Silva , L. Affengruber , R. Spjker , J. Thomas ,
R.C. Berg , Meropi Kontogiani , M. Sousa , C. Kontogiorgis ,
G. Gartlehner , on behalf of working group 3 in the EVBRES COST
Action (https://evbres.eu)

Please cite this article as: B. Nussbaumer-Streit , M. Ellen , I. Klerings , R. Sfetcu , N. Riva , M. Mahmić-Kaknjo , G. Poulentzas , P. Martinez , E. Baladia , L.E. Ziganshina , M.E. Marqués , L. Aguilar , A.P. Kassianos , G. Frampton , A.G. Silva , L. Affengruber , R. Spjker , J. Thomas , R.C. Berg , Meropi Kontogiani , M. Sousa , C. Kontogiorgis , G. Gartlehner , on behalf of working group 3 in the EVBRES COST Action (https://evbres.eu), Resource use during systematic review production varies widely: a scoping review, *Journal of Clinical Epidemiology* (2021), doi: https://doi.org/10.1016/j.jclinepi.2021.05.019

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Highlights

- Evidence on resource use is limited to studies reporting mostly on the resource "time" and not always under real life conditions.

- Administration and project management, study selection, data extraction, and critical appraisal seem to be very resource intensive, varying with the number of included studies, while protocol development, literature search, and study retrieval take less time.

- Lack of experience and domain knowledge, lack of collaborative and supportive software, as well as lack of good communication and management can increase resource use during the systematic review process.

Resource use during systematic review production varies widely: a scoping review

Nussbaumer-Streit B[1]*, Ellen M[2,3], Klerings I[1], Sfetcu R[4,5], Riva N6, Mahmić-Kaknjo M[7,8], Poulentzas G[9], Martinez P[10,11], Baladia E[10], Ziganshina LE[12], Marqués ME[10], Aguilar L[10], Kassianos AP[13,14], Frampton G[15], Silva AG[16], Affengruber L[1,17], Spjker R[18,19], Thomas J[20], Berg RC[21], Meropi Kontogiani[22], Sousa M[23,24], Kontogiorgis C[9], Gartlehner G[1,25], on behalf of working group 3 in the EVBRES COST Action (https://evbres.eu)

1 Cochrane Austria, Danube University Krems, Krems a.d. Donau, Austria

2 Department of Health Systems Management, Guilford Glazer Faculty of Business and Management and Faculty of Health Sciences, Ben-Gurion University of the Negev, Israel

3 Institute of Health Policy Management and Evaluation, Dalla Lana School Of Public Health, University of Toronto, Canada

4 National School of Public Health, Management and Professional Development Bucharest, Romania

5 Spiru Haret University, Faculty of Psychology and Educational Sciences

6 Department of Pathology, Faculty of Medicine and Surgery, University of Malta, Msida, Malta

7 Department of Clinical Pharmacology, Cantonal Hospital Zenica, Zenica, Bosnia and Herzegovina

8 Faculty of Medicine, University of Zenica, Zenica, Bosnia and Herzegovina

9 Laboratory of Hygiene and Environmental Protection, Department of Medicine, Democritus University of Thrace

10 Centro de Análisis de la Evidencia Científica, Academia Española de Nutrición y Dietética, España

11 Techné research group. Department of knowledge engineering of the Faculty of Science. University of Granada. Spain

12 Russian Medical Academy for Continuing Professional Education (RMANPO) and the Kazan State Medical University of the Ministry of Health of Russian Federation

13 Department of Applied Health Research, University College London, London, UK

14 Department of Psychology, University of Cyprus, Nicosia, Cyprus

15 Southampton Health Technology Assessments Centre (SHTAC), Faculty of Medicine, University of Southampton, UK

16 School of Health Sciences & CINTESIS.UA, University of Aveiro, Campus UNiversitário de Santiago, Portugal

17 Department of Family Medicine, Care and Public Health Research Institute (CAPHRI), Maastricht University, The Netherlands

18 Cochrane Netherlands, Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht University, Utrecht, the Netherlands

19 Amsterdam UMC, Univ of Amsterdam, Amsterdam Public Health, Medical Library, Meibergdreef 9, Amsterdam, Netherlands

20 EPPI-Centre, UCL, London, UK

21 Norwegian Institute of Public Health, Oslo, Norway

22 Department of Nutrition and Dietetics, School of Health Sciences and Education, Harokopio University, Athens, Greece

23 Nutrition & Metabolism, NOVA Medical School, Faculdade de Ciências Médicas, Universidade NOVA de Lisboa, Campo dos Mártires da Pátria, 1169-056 Lisboa, Portugal

24 CINTESIS, NOVA Medical School, NMS, Universidade Nova de Lisboa, Campo dos Mártires da Pátria, 1169-056 Lisboa, Portugal

25 RTI International, Research Triangle Park, North Carolina, USA

*Nussbaumer-Streit Barbara (corresponding author)

Cochrane Austria, Danube University Krems

Dr.-Karl-Dorrek Strasse 30,

3500 Krems a.d. Donau, Austria

Email: barbara.nussbaumer-streit@donau-uni.ac.at

 +43 (0)2732 893-2919

Fax: +43 (0)2732-893-4910

Abstract

Objective: We aimed to map the resource use during systematic review (SR) production and reasons why steps of the SR production are resource intensive to discover where the largest gain in improving efficiency might be possible.

Study Design and Setting: We conducted a scoping review. An information specialist searched multiple databases (e.g., Ovid MEDLINE, Scopus) and implemented citation-based and grey literature searching. We employed dual and independent screenings of records at the title/abstract and full-text levels and data extraction.

Results: We included 34 studies. Thirty-two reported on the resource use—mostly time; four described reasons why steps of the review process are resource intensive. Study selection, data extraction, and critical appraisal seem to be very resource intensive, while protocol development, literature search, or study retrieval take less time. Project management and administration required a large proportion of SR production time. Lack of experience, domain knowledge, use of collaborative and SR-tailored software, and good communication and management can be reasons why SR steps are resource intensive.

Conclusion: Resource use during SR production varies widely. Areas with the largest resource use are administration and project management, study selection, data extraction, and critical appraisal of studies.

Keywords: evidence synthesis, time, personnel, costs, resources, efficient

Running title: resource use during systematic review production

What is new

Key findings:

- Overall we identified 34 studies. Of these, 32 reported on resource use: three looked at resource use across the complete SR production process, while others only assessed the resource use of selected steps of the review process. Mostly, studies reported on the resource "time."

- Administration and project management, study selection, data extraction, and critical appraisal seem to be very resource intensive, varying with the number of included studies, while protocol development, literature search, and study retrieval take less time.

- Four studies reported on reasons why steps are resource intensive: lack of experience and domain knowledge, lack of collaborative and SR-tailored software, doing steps manually instead of using supportive software, and lack of good communication and management of the SR process.

What this adds to what is known: Our scoping review is the first to give an overview of the resource use during SR production. It shows that evidence on resource use is limited to studies reporting mostly on the resource "time" and not always under real life conditions. It also reveals that administration/project management is a time-consuming task that should be considered an important part of the SR process.

What is the implication, what should change now: Based on our results, methods and tools to support project management and administration throughout a project, as well as methods and tools to speed up study selection, data extraction, and critical appraisal could help save resources.

1. Introduction

Well-conducted systematic reviews (SRs) are considered the most reliable form of evidence syntheses because they employ high methodological standards in summarizing primary research. SRs play an important role as support for evidence-based clinical and health policy decision-making. However, conducting a SR is very resource intensive and can take up to 2 years for completion (1, 2). This often does not meet the needs of decision-makers, especially in times where evidence syntheses must answer urgent questions such as during the ongoing coronavirus pandemic.

SRs are also essential in primary research. Systematically reviewing the existing evidence before starting a new study is important to ensure high quality and relevant primary research (3). Knowing all the studies in a field helps focus on topics and research questions that require new studies. In addition, learning from earlier studies helps optimally design new studies (4, 5). Ideally, a comprehensive up-to-date SR informs every new study. However, as the methodological standards in SR production are very high, the steps to develop or update a SR are complex and resource intensive. This can keep primary researchers from conducting or updating a SR (6).

According to Cochrane, a SR is "a review of a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies" (7). To conduct a SR, certain steps must be completed, from formulating a focused research question to conducting a comprehensive search strategy, screening the identified literature, and critically appraising and synthesizing the primary studies (8). All these steps require resources such as time or money to complete. Performing a high quality systematic search, for example, requires an information specialist's expertise and time. A recent study showed that information

specialists needed an average aggregated time of 26.9 hours when developing a search strategy (9). Another study assessed the resource need for completing a SR with meta-analyses as ranging from 216 to 2518 hours, depending on the number of studies included and the comparisons and outcomes assessed (10).

The production and update of SRs must become more resource efficient to meet the time-sensitive needs of clinical and health policy decision-makers and to facilitate the uptake of an evidence-based research approach by primary researchers. To discover where the largest gain in improving efficiency might be possible, we wanted to map the resource use of different SR steps. This could help identify the most resource-intensive areas in the review process. In addition, we wanted to know the reasons why certain steps of the review process are resource intensive, in order to identify suitable methods to address them. Therefore, we conducted a scoping review of the published literature mapping the resource use of each step of the SR process and the reasons why steps of the SR process are resource intensive. Specifically, we strove to answer the following two key questions (KQs):

KQ 1. How many resources (e.g., time, costs) do different steps of the SR production consume?

KQ 2. What are the reasons why some steps of the SR production are resource intensive?

2. Materials and Methods

This scoping review is part of working group 3's work within the EVBRES (EVidence-Based RESearch) – COST Action CA17117, funded by the European Union (www.evbres.eu). The protocol for this scoping review was published a priori via the Open Science Framework (https://osf.io/fby54/) (11).

2.1. Study design

We conducted a scoping review. Within EVBRES, we define a scoping review as a form of knowledge synthesis that addresses an exploratory research question aimed at mapping key concepts, theories, types and sources of evidence, and gaps in research related to a defined area or field by systematically searching, selecting, and synthesizing existing knowledge. We followed the guidance given by Arksey and O'Malley (12), Levac et al., and Peters et al. (13, 14). We reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Extension for Scoping Reviews (PRISMA-ScR) (15).

2.2. Information sources and search strategies

The search for this scoping review followed the three-step process recommended by the Joanna Briggs Institute (16):

1) In a first step, an information specialist (IK) conducted a focused search of Ovid MEDLINE and Current Contents Connect (Web of Science) in November 2019. We screened these results to identify relevant studies for inclusion. Using PubReMiner and AntConc (17, 18), the included studies were analyzed to identify relevant text words contained in the title and abstract as well as Medical Subject Headings (MeSH) terms.

2) Based on search terms derived from these included studies, IK developed a second, more comprehensive search strategy and searched the following databases in May 2020: Ovid MEDLINE, Scopus (Elsevier), Science Citation Index Expanded, Social Sciences Citation Index, and Current Contents Connect (all via Web of Science). The Ovid MEDLINE strategy

was reviewed by a second information specialist (RS) in accordance with the Peer Review of Electronic Search Strategies (PRESS) guideline (19).

3) Using the studies identified by these searches, we conducted citation-based searches: manual screening of reference lists, forward citation tracking (via Scopus in May 2020), and a similar articles search (via PubMed, limited to the first 50 linked references for each seed article, in May 2020).

The database searches were limited to studies published since 2009 (the year when PRISMA was published). Last, to identify grey literature, we contacted experts in the field and screened the latest proceedings of the Cochrane Colloquium (20) to also include new studies that might not have been published yet. Details on the search strategies are available in web appendix 1.
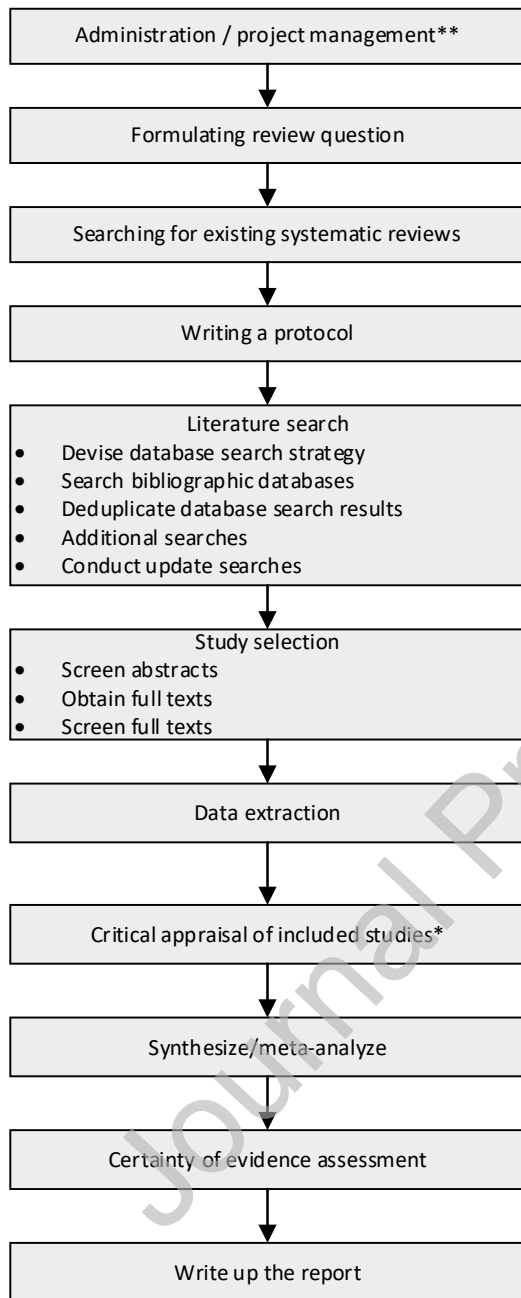
2.3. Eligibility criteria

The eligibility criteria are specified in Table 1. For KQ 1, we included studies that assessed the resource needs of one or more steps of the review process. Studies that tracked resource use when conducting SRs and those that asked for resource use via a survey as well as studies that modeled resource use were eligible. If a study only reported relative measures of resource use, such as time or workload saved, without specifying the absolute resource use of the task of interest, the study was excluded. To define the SR production steps, we used the list of steps provided by Tsafnat et al. 2014 (8) and added other important steps: "critical appraisal" and "grade the certainty of evidence," (21). When mapping the results, we identified an additional SR production step and added "Administration/project management" a posteriori (see Figure 1). For KQ 2, we were interested in qualitative studies that asked SR authors via interviews or surveys why some steps of the SR production are resource intensive. We limited the publication date to 2009 because the current reporting standards, the PRISMA guidelines, were published at that time.

Table 1: Eligibility criteria

	Inclusion criteria	**Exclusion criteria**
Topic	KQ1: Studies reporting the resource use of producing or updating SRs* KQ2: Studies reporting reasons why steps of the review production are resource intensive	Studies not reporting resource use
Concept	KQ1 and KQ2: Addressing the resource use of one or more steps of a SR (as depicted in Figure 1) of health interventions or diagnostic or prognostic studies	Studies assessing other steps such as the dissemination of a SR's results
Outcomes	KQ1: • Time spent on tasks (e.g., minutes per task) • Effort of personnel (e.g., full-time equivalent) • Costs (e.g., salary, license fees) • Material used KQ2: • Perceived reasons why some steps are/ are not resource intensive (barriers/ facilitators) • Perceived resource-intensive steps	Any other outcomes, such as time or money saved, without specification of the total resource amount spent on a task
Study design/ publication type	KQ1: • SRs • Empirical studies: measuring resource use by tracking • Surveys: asking about resource use • Simulation studies: modeling resource use KQ2: • Qualitative studies (e.g., interviews, open surveys)	Other publications (e.g., editorials, letters)
Full text availability		If full text was not retrievable via our libraries
Timing	2009–2020	
Language	All languages	
Abbreviations: KQ: key question, SR: systematic review **resource use must already be mentioned in the abstract*		

Figure 1: Steps of a systematic review

```
┌─────────────────────────────────────────────┐
│   Administration / project management**      │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│        Formulating review question          │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│    Searching for existing systematic reviews │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│             Writing a protocol               │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│             Literature search                │
│  • Devise database search strategy           │
│  • Search bibliographic databases            │
│  • Deduplicate database search results       │
│  • Additional searches                       │
│  • Conduct update searches                   │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│              Study selection                 │
│  • Screen abstracts                          │
│  • Obtain full texts                         │
│  • Screen full texts                         │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│              Data extraction                 │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│     Critical appraisal of included studies*  │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│           Synthesize/meta-analyze            │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│        Certainty of evidence assessment      │
└─────────────────────────────────────────────┘
                      │
                      ▼
┌─────────────────────────────────────────────┐
│             Write up the report              │
└─────────────────────────────────────────────┘
```

*added a priori to Tsafnat et al.'s list (8)

** added after synthesis

2.4. Selection of sources of evidence

After piloting the abstract screening with 50 records, the author team used Covidence

(www.covidence.org) to dually and independently screen records based on titles and abstracts

and full texts. We resolved any screening discrepancies at the abstract or full-text level by discussion or by consultation with a third author. We stored records in the reference management software EndNote (22).

2.5. Data charting process and data items

We developed and piloted a data extraction form using Google Forms. Two team members independently extracted data; a third author made final decisions in cases of discrepancies.

2.6. Data synthesis

We mapped the results of the scoping review in a summary table, grouping them along the steps of the review process. Because the goal of this scoping review was to descriptively map the resource use of SR production steps rather than to derive cause–effect relationships, we did not apply a formal certainty of evidence or risk of bias (RoB) assessment.

3. Results

We included 34 studies (32 quantitative primary studies, 1 qualitative primary study, 1 SR) that were published in 38 publications (9, 23-59) (see Figure 2: PRISMA study flow). In the following sections, we first summarize the characteristics of the included studies. We then present the resource use for single steps of the review process as well as across all the steps combined. Finally, we summarize the reasons why the review authors perceive some steps of the SR process as resource intensive.

Figure 2: PRISMA study flow
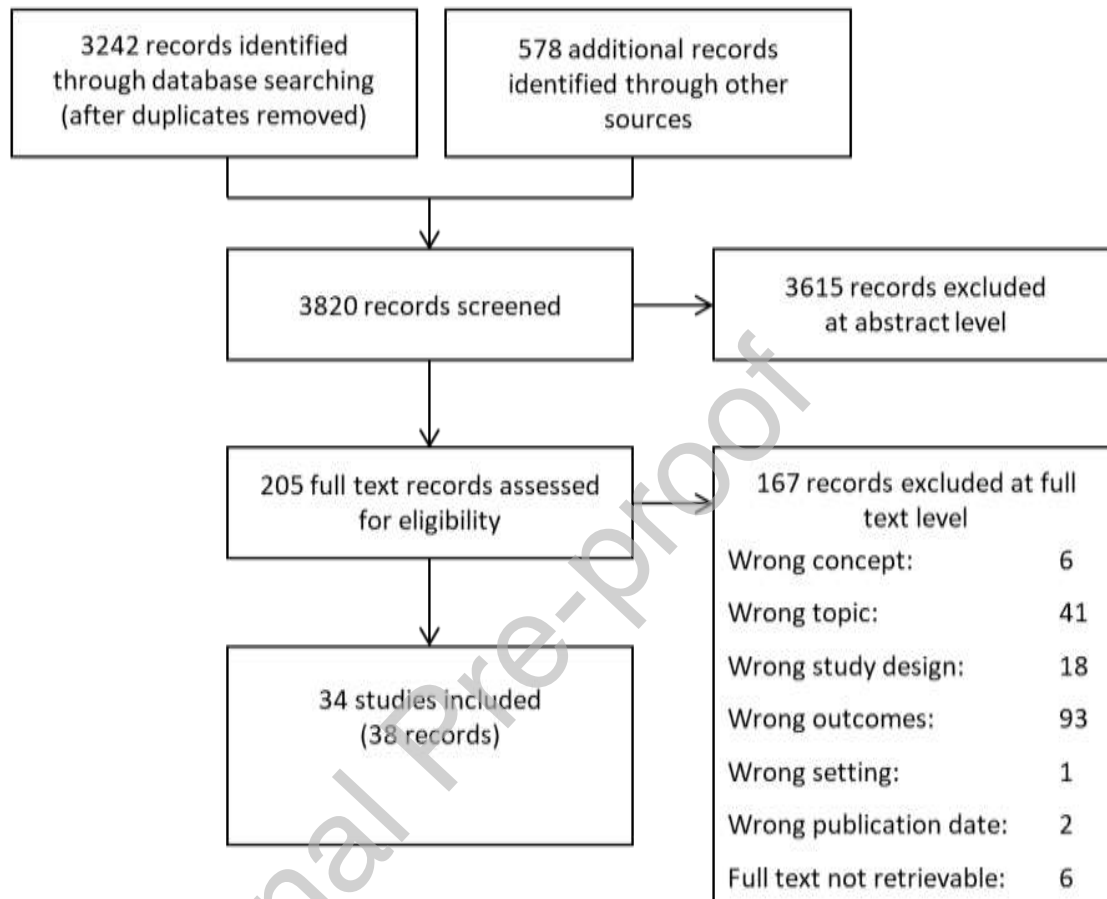
3.1. Characteristics of the included studies

Thirty-two studies contributed to KQ1 and reported on resource use, mostly on time spent (9, 23-39, 41-46, 48-59). Of these, three studies compared resource use across all steps of the review process (28, 37, 52), while the other 29 focused on one or more single steps: study selection (title/abstract screening n = 7, obtaining full-text articles n = 1, full-text screening n = 4), literature search (n = 11), data extraction (n = 7), and critical appraisal (n = 4). Four studies helped answer KQ2 and described reasons why certain steps of the review process are resource intensive (23-25, 28, 40, 47). In web appendix 2, we summarize the characteristics of all the included studies. In web appendix 3, we map the resource use to the steps of the SR process.

3.2. Comparative resource use across all steps of the systematic review process

Three studies assessed the resource use required to conduct complete SRs (28, 37, 52). Haddaway et al. (37) and Pham et al. (52) simulated SR production and estimated the resource needs for each step based on survey responses from SR authors or on the literature. According to Pham et al., the mean time needed to conduct all SR tasks was 881 person-hours, ranging from 243 to 1752 person-hours, depending on the scope of the SR. When comparing different steps of the review process, protocol development and literature search were less resource intensive than selecting studies and extracting data, which together took roughly 50 % of all the time resources needed (52). Haddaway et al. simulated that the average time to conduct a SR is 1312 person-hours. They also concluded that study selection as well as data extraction and critical appraisal of studies are very time consuming. Searching, retrieving full texts, and assembling a library of relevant studies required less time than most stages. The largest proportion of time, however, was needed for administration and planning (37). Clark et al. (28) published the only empirical study tracking the time needed to conduct one SR with the help of automation tools. Their SR identified 1694 abstracts and finally included 8 publications. They completed it in 61 person-hours. The least time-consuming tasks were searches, deduplication, and obtaining full texts. The most time-consuming steps were administration and project management, abstract and full-text screening, data extraction, critical appraisal, and data synthesis (28).

3.3. Resource use during literature search

The median time for the whole literature search process ranged from 7.85 to 22 hours (9, 54). Studies also reported the resource use for different literature search components. Developing a search strategy for a SR took 1 to 13.5 hours, depending on the software support (text mining) used (9, 26, 28, 35, 39). Translating a search strategy for another database took 11 to 79 minutes manually and 6 to 57 min per database using Polyglot Search Translator (29). Deduplication of records was estimated to take 3 to 10 min per SR, depending on the software

used; however, the authors also report deduplicates missed by the software that required manual reduplicating later in the process (45).

Studies reported the resource use of different additional search approaches: manual reference list checking of included studies can take 8 hours compared to 3 hours when using Scopus (27). Hand searching, the manual examination of content from relevant journals or conference proceedings, ranged from 6 to 60 min per source (30). Contacting authors took nearly 7 hours in a study that contacted 88 authors (31). Citation chasing required about 1 hour (30), and web searching added from 8 hours (30) to 3 days (59). The median time to search grey literature took 85 minutes (range: 20 to 3480) (54).

3.4. Resource use during study selection

Team members were able to screen from 0.13 to 2.88 abstracts per minute (32-34, 56). Conflict resolution–often necessary in dual screening processes–took on average 5 min per conflict, and retrieving full-text articles took 4 minutes per full text (56).

Different approaches to title/abstract screening exist, and although they can reduce the time needed for this step, they may also increase the full-text screening burden. Shemilt et al. (56) analyzed the following approaches: (1) single screening; (2) single screening with semiautomated software (one person screens abstracts and automatically excludes those deemed irrelevant by the algorithm); (3) dual screening (two people independently screen all abstracts and discuss conflicting decisions); and (4) safety first (two people independently screen all abstracts, all potentially relevant abstracts are included for full-text screening). These approaches were all followed by dual independent full-text screening. Based on a project with 12477 abstracts, the semiautomated approach took the least time (572 hours) and cost the least (£ 37,860), while dual screening took twice as long (1089 hours) and cost nearly twice as much (£ 75139) (56).

Two studies assessed a type of crowdsourcing for title/abstract screening. While experts cost between $ 3034 and $ 8777 to complete such a task, the crowd costs ranged between $ 458 and $ 2223. However, it required the crowd 4 to 17 days to complete the screening (49). In another study, the median time to acquire enough assessments per citation was 42 days (50). For acquiring enough assessments for all full texts, the crowd took another median of 36 days (50).

Full-text screening took people from 4.3 to 5 minutes per full text (34, 56). Conflict resolution took 5 minutes per full text (56).

One study showed that diagnostic test accuracy (DTA) reviews have on average 185 % more workload during abstract screening and 167 % more during full-text screening, because searches for DTA reviews identify many more records (51).

3.5. Resource use during data extraction

Extracting major information on study design, participants, and results took one person an average of 41 to 65 min per study (36, 57). Using two monitors instead of one helped reduce the time spent on data extraction (57); experience in data extraction was also associated with less time spent on this task (41). While single data abstraction and verification took on average 107 minutes per study, doing dual independent data abstraction took 172 min (46). Data extraction from trial registry entries took on average 40 minutes per study (53), while manual data extraction took on average 11 to 13 minutes per figure; using software reduced the time spent to 5–6 minutes per figure (42).

Sometimes the SR team must translate a study into English before being able to extract data. Using Google Translate took on average 15 to 60 minutes per study, depending on the publication's original language (Spanish: 15 min to Chinese: 60 min)(23, 24).

3.6. Resource use during critical appraisal

On average, one person needed 21 minutes to apply the Cochrane RoB Tool 1.0 per RCT (38), 9.5 minutes to apply RoBANS, and 10.45 minutes for MINORS (44). Applying ROBINS-I plus finding consensus took the longest, at 48.45 minutes, while applying ROB-NRSE and reaching consensus took on average 36.98 minutes (43). Authors needed 10 to 15 minutes to critically appraise a SR using the AMSTAR tool (55).

3.7. Reasons for perceived resource intensity

Facilitators: Methodological experience and content knowledge of team members, existing data extraction sheets from former projects, blocking time for the SR, daily project meetings to discuss upcoming questions, and writing the protocol in past tense all contributed to speeding up the process. In addition, team members' physical proximity allowed for ongoing communication and short time lapses between tasks, and familiarity with the used software tools helped (28).

Barriers: Lack of domain expertise, juggling other projects with competing deadlines, noisy surroundings, resource unavailability, poor internet, and software incompatibilities and limitations (e.g., automation of only one task) increased the resource use (28). Lack of collaborative platforms, of literature search experience (47), and manually doing tasks such as abstract screening acted as further barriers (40). Hoang et al. also identified a gap between the available SR support tools and the tools used. The majority of review authors still rely only on Excel and EndNote, not using other SR-tailored tools. Potential explanations were that the reviewers are not aware of the other tools, might have limited access due to costs, or feel more comfortable using well-known tools (40). Balk et al. highlighted that automated translation needs more resources for non-European languages and tables than for European languages, and if manuscripts are not available in Word or PDF formats (23).

4. Discussion

To the best of our knowledge, this is the first scoping review mapping the resource use required to conduct a SR. Across all SR production steps, study selection, data extraction, and critical appraisal seem to be very resource intensive while protocol development, literature search, and study retrieval take less time (28, 37, 52). Project management and coordination needed the largest proportion of SR production time. This is relevant for future initiatives that aim to make the SR production process more efficient, since this task is usually not in focus when thinking about SR production steps.

We did not identify any study reporting on how much resources the certainty of evidence assessments require, and only 6 studies reported resources other than time. We also did not identify a study focusing specifically on updating an existing SR; however, the resource use for individual steps of the SR process is probably similar when updating a review. The only study reporting resource use for the specific steps of formulating the review question, searching for existing SRs, writing a protocol, synthesis/ meta-analysis, and writing up the report was that of Clark et al. (28). This study was a case study of a single SR, so the generalizability is very limited.

Our scoping review showed that literature search requires only a small proportion of the overall time required to complete a SR. This contradicts survey results among scientists in dentistry who perceived literature search as a particularly time-consuming challenge (47). In the included studies of our scoping review, information specialists performed the search steps, which might explain why literature search did not take so much time though it is a very complex step in the SR process.

The time needed for study selection, data extraction, and critical appraisal varied largely depending on the number of identified records and full texts. Other factors such as lack of experience or domain expertise might also increase resource use. In addition, the lack of using

collaborative and supportive software increased the resource need (28, 40). Although many supportive tools are available, the use of automation tools is still not very common in the SR community (60). Clark 2020a highlighted that the least time-consuming SR production tasks were generally those where the most automation tools were available, and vice versa (28).

4.1. Limitations

Our scoping review has several limitations. First, we focused only on the resource use of different steps and approaches, not on the validity and accuracy of the methods or tools. We plan to answer this in a follow-up scoping review (https://osf.io/9423z). Second, we did not assess the RoB of the included studies. Instead, we described the methodological approach of the included studies for transparency. Third, we limited the inclusion criteria to studies published from 2009 onward—the year when PRISMA was published. We might have missed older but informative and relevant studies. However, we think that the included studies provide a good overview of current resource use in SR production. Fourth, we included only studies that mention any type of resource use in the abstract. We considered this necessary to achieve a balance of sensitivity and specificity during the screening process. Fifth, the included studies are heterogeneous. Specific time estimates must be interpreted with caution since they also depend on contextual factors such as the topic of the SR or team characteristics. However, we think that the median or mean estimates as well as the ranges provide a good orientation.

5. Conclusion

Evidence on resource use during SR production is limited to studies reporting mostly on the resource "time" - often not under real life conditions. To be able to gain a more comprehensive understanding of the resource requirements for SRs, future studies need to assess resource use prospectively across various types of reviews (e.g. intervention, DTA, prognosis) and across different teams and settings to better reach generalizable estimates. Based on the identified

evidence, the areas with the largest resource use are administration and project management, study selection, data extraction, and critical appraisal of studies.

6. Acknowledgments

7. Funding Source

8. Conflict of Interest Disclosures

None of the authors report any financial conflicts of interest with respect to the topic of this manuscript. All authors have a general interest in evidence synthesis methods. Some authors are associated with groups, conferences, and tools focusing on evidence synthesis methods:

Barbara Nussbaumer-Streit and Gerald Gartlehner are co-convenors, Lisa Affengruber associate co-convenor of the Cochrane Rapid Reviews Methods Group.

Raluca Sfetcu is a member of the JBI method group for "Systematic reviews of etiology and risk".

Patricia Martinez is associated with the Techné Co-word research group that has patented knowledge engineering software that is used for mapping.

Rene Spijker is on the organizing committee of ICASR, member of the Cochrane Information Specialist Executive, and Editorial Board Member of BMJ Evidence Based Medicine.

James Thomas is on the organizing committee of ICASR (International Collaboration for the automation of systematic reviews); a co-senior scientific editor for the Cochrane Handbook;

leads / contributes to research projects on systematic review automation; and leads the team that develops and supports EPPI-Reviewer, a software platform for evidence synthesis.

Author statements

Nussbaumer-Streit B: conceptualization, methodology, validation, formal analysis, investigation, data curation, writing – original draft, project administration

Ellen M: conceptualization, methodology, investigation, writing- review & editing

Klerings: conceptualization, literature search, writing-review & editing

Sfetcu R: conceptualization, methodology, investigation, writing- review & editing

Riva N: E: conceptualization, methodology, investigation, writing- review & editing

Mahmić-Kaknjo M: conceptualization, methodology, investigation, writing- review & editing

Poulentzas G: investigation, writing- review & editing

Martinez P: E: conceptualization, methodology, investigation, writing- review & editing

Baladia E: conceptualization, methodology, investigation, writing- review & editing

Ziganshina LE: E: conceptualization, methodology, investigation, writing- review & editing

Marqués ME: E: conceptualization, methodology, investigation, writing- review & editing

Aguilar L: E: conceptualization, methodology, investigation, writing- review & editing

Kassianos AP: E: conceptualization, methodology, investigation, writing- review & editing

Frampton G: conceptualization, methodology, investigation, writing- review & editing

Silva AG: E: conceptualization, methodology, investigation, writing- review & editing

Affengruber L: E: conceptualization, methodology, investigation, writing- review & editing

Spjker R: conceptualization, methodology, writing- review & editing

Thomas J: conceptualization, methodology, writing- review & editing

Berg RC: conceptualization, methodology, writing- review & editing

Kontogiani M: conceptualization, methodology, investigation, writing- review & editing

Sousa M: conceptualization, methodology, writing- review & editing

Kontogiorgis C: investigation, writing- review & editing

Gartlehner G: conceptualization, methodology, writing- review & editing

References

1. Ganann R, Ciliska D, Thomas H. Expediting systematic reviews: methods and implications of rapid reviews. Implement Sci. 2010;5:56.

2. Hartling L, Guise JM, Kato E, Anderson J, Aronson N, Belinson S, et al. Agency for Healthcare Research and Quality (US). 2015:02.

3. Clarke M, Hopewell S, Chalmers I. Clinical trials should begin and end with systematic reviews of relevant evidence: 12 years and waiting. The Lancet. 2010;376(9734):20-1.

4. Robinson KA, Brunnhuber K, Ciliska D, Juhl CB, Christensen R, Lund H. Evidence-Based Research Series-Paper 1: What Evidence-Based Research is and why is it important? J Clin Epidemiol. 2021;129:151-7.

5. Lund H, Juhl CB, Nørgaard B, Draborg E, Henriksen M, Andreasen J, et al. Evidence-Based Research Series-Paper 2 : Using an Evidence-Based Research approach before a new study is conducted to ensure value. J Clin Epidemiol. 2021;129:158-66.

6. Clayton GL, Smith IL, Higgins JPT, Mihaylova B, Thorpe B, Cicero R, et al. The INVEST project: investigating the use of evidence synthesis in the design and analysis of clinical trials. Trials. 2017;18(1):219.

7. Cochrane Community. Glossary [Available from: https://cdev.cochrane.org/glossary#letter-S.

8. Tsafnat G, Glasziou P, Choong MK, Dunn A, Galgani F, Coiera E. Systematic review automation technologies. Syst. 2014;3:74.

9. Bullers K, Howard AM, Hanson A, Kearns WD, Orriola JJ, Polo RL, et al. It takes longer than you think: librarian time spent on systematic review tasks. J Med Libr Assoc. 2018;106(2):198-207.

10. Allen IE, Olkin I. Estimating time to conduct a meta-analysis from number of citations retrieved. Jama. 1999;282(7):634-5.

11. Nussbaumer-Streit B, Ellen M, Klerings I, Gartlehner G, Thomas J, Mikkelsen LR, et al. Identifying resource intensive areas of systematic review production and updating – a scoping review 2020 [Available from: https://osf.io/8an4j.

12. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8(1):19-32.

13. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci. 2010;5(1):69.

14. Peters MDJ, Marnie C, Tricco AC, Pollock D, Munn Z, Alexander L, et al. Updated methodological guidance for the conduct of scoping reviews. JBI Evid Synth. 2020;18(10).

15. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med. 2018;169(7):467-73.

16. Peters MDJ, Godfrey C, McInerney P, Munn Z, Tricco AC, Khalil H. Chapter 11: Scoping Reviews (2020 version). In: Aromataris E, Munn Z (Editors). JBI Manual for Evidence Synthesis, JBI, 2020 2020 [Available from: https://wiki.jbi.global/display/MANUAL/Chapter+11%3A+Scoping+reviews.

17. Koster JA. PubReMiner 2014 [Available from: https://hgserver2.amc.nl/cgi-bin/miner/miner2.cgi.

18. Anthony L. AntConc (Version 3.5.9) [Computer Software]. Tokyo, Japan: Waseda University; 2020.

19. 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.

20. The Cochrane Collaboration. Advances in Evidence Synthesis: special issue Cochrane Database of Systematic Reviews 2020;(9 Suppl 1) 2020 [Available from: https://doi.org/10.1002/14651858.CD202001.

21. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021): Cochrane; 2021 [Available from: www.training.cochrane.org/handbook].

22. Clarivate. EndNote X8.

23. Balk EM, Chung M, Chen ML, Chang LK, Trikalinos TA. Data extraction from machine-translated versus original language randomized trial reports: a comparative study. Syst. 2013;2:97.

24. Balk EM, Chung M, Chen ML, Trikalinos TA, Kong Win Chang L. Assessing the Accuracy of Google Translate to Allow Data Extraction From Trials Published in Non-English Languages. AHRQ Methods for Effective Health Care. Rockville (MD): Agency for Healthcare Research and Quality (US); 2013.

25. Balk EM, Chung M, Hadar N, Patel K, Yu WW, Trikalinos TA, et al. AHRQ Methods for Effective Health Care. Accuracy of Data Extraction of Non-English Language Trials with Google Translate. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012.

26. Bramer WM, Rethlefsen ML, Mast F, Kleijnen J. Evaluation of a new method for librarian-mediated literature searches for systematic reviews. Res. 2018;9(4):510-20.

27. Chapman AL, Morgan LC, Gartlehner G. Semi-automating the manual literature search for systematic reviews increases efficiency. Health Info Libr J. 2010;27(1):22-7.

28. Clark J, Glasziou P, Del Mar C, Bannach-Brown A, Stehlik P, Scott AM. A full systematic review was completed in 2 weeks using automation tools: a case study. J Clin Epidemiol. 2020;121:81-90.

29. Clark JM, Sanders S, Carter M, Honeyman D, Cleo G, Auld Y, et al. Improving the translation of search strategies using the polyglot search translator: A randomized controlled trial. J Med Libr Assoc. 2020;108(2):195-207.

30. Cooper C, Booth A, Britten N, Garside R. A comparison of results of empirical studies of supplementary search techniques and recommendations in review methodology handbooks: a methodological review. Syst. 2017;6(1):234.

31. Cooper C, Bou JT, Varley-Campbell J. Evaluating the effectiveness, efficiency, cost and value of contacting study authors in a systematic review: a case study and worked example. BMC Med Res Methodol. 2019;19(1):45.

32. Felizardo KR, Barbosa EF, Maldonado JC, editors. A visual approach to validate the selection review of primary studies in systematic reviews: A replication study. Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE; 2013.

33. Felizardo KR, Salleh N, Martins RM, Mendes E, Macdonell SG, Maldonado JC, editors. Using visual text mining to support the study selection activity in systematic

literature reviews. International Symposium on Empirical Software Engineering and Measurement; 2011.

34. Giummarra MJ, Lau G, Gabbe BJ. Evaluation of text mining to reduce screening workload for injury-focused systematic reviews. Inj Prev. 2019;26:26.

35. Grames EM, Stillman AN, Tingley MW, Elphick CS. An automated approach to identifying search terms for systematic reviews using keyword co-occurrence networks. Methods Ecol Evol. 2019:NIL_1-NIL_10.

36. Gresham G, Matsumura S, Li T. Faster may not be better: data abstraction for systematic reviews. Cochrane Colloquium; Hyderabad. AR I2014.

37. Haddaway NR, Westgate MJ. Predicting the time needed for environmental systematic reviews and systematic maps. Conserv. 2019;33(2):434-43.

38. Hartling L, Bond K, Vandermeer B, Seida J, Dryden DM, Rowe BH. Applying the Risk of Bias Tool in a Systematic Review of Combination Long-Acting Beta-Agonists and Inhaled Corticosteroids for Persistent Asthma. PLoS ONE. 2011;6(2):6.

39. Hausner E, Guddat C, Hermanns T, Lampert U, Waffenschmidt S. Development of search strategies for systematic reviews: validation showed the noninferiority of the objective approach. J Clin Epidemiol. 2015;68(2):191-9.

40. Hoang L, Schneider J. Opportunities for computer support for systematic reviewing - a gap analysis. Transform Digit Worlds (2018). 2018;10766:367-77.

41. Horton J, Vandermeer B, Hartling L, Tjosvold L, Klassen TP, Buscemi N. Systematic review data extraction: cross-sectional study showed that experience did not increase accuracy. J Clin Epidemiol. 2010;63(3):289-98.

42. Jelicic Kadic A, Vucic K, Dosenovic S, Sapunar D, Puljak L. Extracting data from figures with software was faster, with higher interrater reliability than manual extraction. J Clin Epidemiol. 2016;74:119-23.

43. Jeyaraman MM, Rabbani R, Copstein L, Robson RC, Al-Yousif N, Pollock M, et al. Methodologically rigorous risk of bias tools for nonrandomized studies had low reliability and high evaluator burden. J Clin Epidemiol. 2020;128:140-7.

44. Kim SY, Park JE, Lee YJ, Seo HJ, Sheen SS, Hahn S, et al. Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. J Clin Epidemiol. 2013;66(4):408-14.

45. Kwon Y, Lemieux M, McTavish J, Wathen N. Identifying and removing duplicate records from systematic review searches. J Med Libr Assoc. 2015;103(4):184-8.

46. Li T, Saldanha IJ, Jap J, Smith BT, Canner J, Hutfless SM, et al. A randomized trial provided new evidence on the accuracy and efficiency of traditional vs. electronically annotated abstraction approaches in systematic reviews. J Clin Epidemiol. 2019;115:77-89.

47. Major MP, Warren S, Flores-Mir C. Survey of systematic review authors in dentistry: challenges in methodology and reporting. J Dent Educ. 2009;73(4):471-82.

48. Mathes T, Klasen P, Pieper D. Frequency of data extraction errors and methods to increase data extraction quality: a methodological review. BMC Med Res Methodol. 2017;17(1):152.

49. Mortensen ML, Adam GP, Trikalinos TA, Kraska T, Wallace BC. An exploration of crowdsourcing citation screening for systematic reviews. Res. 2017;8(3):366-86.

50. Nama N, Sampson M, Barrowman N, Sandarage R, Menon K, Macartney G, et al. Crowdsourcing the Citation Screening Process for Systematic Reviews: Validation Study. J Med Internet Res. 2019;21(4):e12953.

51. Petersen H, Poon J, Poon SK, Loy C. Increased workload for systematic review literature searches of diagnostic tests compared with treatments: challenges and opportunities. JMIR Med Inform. 2014;2(1):e11.

52. Pham B, Bagheri E, Rios P, Pourmasoumi A, Robson RC, Hwee J, et al. Improving the conduct of systematic reviews: a process mining perspective. J Clin Epidemiol. 2018;103:101-11.

53. Pradhan R, Hoaglin DC, Cornell M, Liu W, Wang V, Yu H. Automatic extraction of quantitative data from ClinicalTrials.gov to conduct meta-analyses. J Clin Epidemiol. 2019;105:92-100.

54. Saleh AA, Ratajeski MA, Bertolet M. Grey Literature Searching for Health Sciences Systematic Reviews: A Prospective Study of Time Spent and Resources Utilized. Evid Based Libr Inf Pract. 2014;9(3):28-50.

55. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. J Clin Epidemiol. 2009;62(10):1013-20.

56. Shemilt I, Khan N, Park S, Thomas J. Use of cost-effectiveness analysis to compare the efficiency of study identification methods in systematic reviews. Syst. 2016;5(1):140.

57. Wang Z, Asi N, Elraiyah TA, Abu Dabrh AM, Undavalli C, Glasziou P, et al. Dual computer monitors to increase efficiency of conducting systematic reviews. J Clin Epidemiol. 2014;67(12):1353-7.

58. Williamson PO. Librarians' reported systematic review completion time ranges between 2 and 219 total hours with most variance due to information processing and instruction. Evidence Based Library and Information Practice. 2019;14(1):80-3.

59. Wright K, Golder S, Rodriguez-Lopez R. Citation searching: a systematic review case study of multiple risk behaviour interventions. BMC Med Res Methodol. 2014;14:73.

60. van Altena AJ, Spijker R, Olabarriaga SD. Usage of automation tools in systematic reviews. Res. 2019;10(1):72-82.