Journal of Clinical Epidemiology 121 (2020) 81−90

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

A full systematic review was completed in 2 weeks using automation tools: a case study

Justin Clark*, Paul Glasziou, Chris Del Mar, Alexandra Bannach-Brown, Paulina Stehlik, Anna Mae Scott

Institute for Evidence-Based Healthcare, Bond University, Gold Coast, Australia

Accepted 18 January 2020; Published online 28 January 2020

Abstract

Background and Objectives: Systematic reviews (SRs) are time and resource intensive, requiring approximately 1 year from protocol registration to submission for publication. Our aim was to describe the process, facilitators, and barriers to completing the first 2-week full SR.

Study Design and Setting: We systematically reviewed evidence of the impact of increased fluid intake, on urinary tract infection (UTI) recurrence, in individuals at risk for UTIs. The review was conducted by experienced systematic reviewers with complementary skills (two researcher clinicians, an information specialist, and an epidemiologist), using Systematic Review Automation tools, and blocked off time for the duration of the project. The outcomes were time to complete the SR, time to complete individual SR tasks, facilitators and barriers to progress, and peer reviewer feedback on the SR manuscript. Times to completion were analyzed quantitatively (minutes and calendar days); facilitators and barriers were mapped onto the Theoretical Domains Framework; and peer reviewer feedback was analyzed quantitatively and narratively.

Results: The SR was completed in 61 person-hours (9 workdays; 12 calendar days); accepted version of the manuscript required 71 person-hours. Individual SR tasks ranged from 16 person-minutes (deduplication of search results) to 461 person-minutes (data extraction). The least time-consuming SR tasks were obtaining full-texts, searches, citation analysis, data synthesis, and deduplication. The most time-consuming tasks were data extraction, write-up, abstract screening, full-text screening, and risk of bias. Facilitators and barriers mapped onto the following domains: knowledge; skills; memory, attention, and decision process; environmental context and resources; and technology and infrastructure. Two sets of peer reviewer feedback were received on the manuscript: the first included 34 comments requesting changes, 17 changes were made, requiring 173 person-minutes; the second requested 13 changes, and eight were made, requiring 121 person-minutes.

Conclusion: A small and experienced systematic reviewer team using Systematic Review Automation tools who have protected time to focus solely on the SR can complete a moderately sized SR in 2 weeks.

Keywords: Systematic reviews; Automation; Methods improvement; 2 week systematic review; 2wSR; Systematic review accelerator; Barriers; Facilitators

1. Background

Systematic reviews (SRs) synthesize evidence to answer a specific question, using methods that are transparent and reproducible. They are considered the highest-level of evidence to underpin clinical and policy decisions.

However, SRs are time and resource intensive, requiring a median of five researchers and 41 weeks to submit to a journal [personal communication, Kathryn Kaiser; [1]]. A median-sized SR search yields 1,781 references (range: 27−92,020) and requires a title/abstract screen of 1,286 references (range 14−77,910) and a full-text screen of 63 references (range: 0−4385). A median-sized SR includes 15 studies (range: 0−291) [1].

The International Collaboration for the Automation of Systematic Reviews was formed to bring together groups working in Systematic Review Automation (SRA), with an aim to enhance the speed, efficiency, and accuracy of SRs [2]. SRA tools remain underused because of lack of social acceptance of the tools, lack of knowledge about their existence, and steep learning curves [3,4].

This study was therefore designed to:

Funding: The systematic review described in the manuscript was conducted as part of the work of the Centre of Research Excellence in Minimising Antibiotic Resistance in the Community (CRE-MARC), funded by the National Health and Medical Research Council (NHMRC), Australia. The funder had no involvement in this systematic review. The present case study was not specifically funded. J.C., P.G., and C.D.M. were involved in the development of systematic review automation tools used in the systematic review.

Conflict of interest: None.

* Corresponding author. Institute for Evidence-Based Healthcare Bond University, 14 University Drive, Robina, Queensland 4226, Australia. Tel.: +61 7 5595 5545.

E-mail address: jclark@bond.edu.au (J. Clark).

https://doi.org/10.1016/j.jclinepi.2020.01.008
0895-4356/

What is new?

Key findings

- Our study shows that it is possible to complete a moderately-sized full systematic review in 2 weeks (2wSR).

What this adds to what was known?

- Systematic reviews currently require approximately one year from protocol registration to submission for publication.
- A small team of experienced systematic reviewers with complementary skills, protected time, and using automation tools, can considerably decrease the current time to completion of a full systematic review.
- We describe the processes, facilitators and barriers to completing a 2wSR.

What is the implication and what should change now?

- High priority systematic reviews can be completed to considerably faster timelines.
- Parallel advancement in time from submission to publication, would result in faster dissemination of knowledge and assist policy uptake.

1) Test whether it is feasible for an experienced SR team to complete a medium-sized SR in 2 weeks, by using SRA tools and blocking off time from other projects;
2) Identify the barriers and facilitators to completing a full SR in 2 weeks (2wSR).

2. Methods

2.1. The systematic review

We completed an SR of randomized controlled trials (RCTs) assessing the impact of increased fluid intake on urinary tract infection (UTI) recurrence, antimicrobial use, and UTI symptoms, in individuals at risk for UTIs [5]. We included RCTs which compared interventions involving increased fluid intake (e.g., water, juice) to those not involving increased fluids. Searches identified 1,694 references; eight trials were included, and 4 meta-analyses were conducted (Table 1). This was a full SR, which followed standard Cochrane methods, is compliant with the Methodological Expectations of Cochrane Intervention Reviews (MECIR), and is reported using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [6,7]. None of the methods used were abbreviated, nor did we use methods adopted in 'rapid reviews.'

2.2. Modifications to standard SR processes

The key modifications to standard SR processes included:

1) Limiting the team to four experienced systematic reviewers with complementary areas of specialization (two clinician researchers, an information specialist, and an epidemiologist).
2) Use of SRA tools (see Tables 2 and 3).
3) Blocked off time from other projects for the duration of the SR ("protected time").
4) A daily meeting to identify the facilitators and barriers to completing the SR and resolve issues as they arise.

An additional modification from our usual processes involved blinding to the Population, Intervention, Comparison and Outcome (PICO) of the SR question. The clinician researchers (PG, CDM) jointly determined the PICO question; however, systematic reviewers conducting the searches, screening, and data extractions (J.C. and A.M.S.) were blinded to the PICO until the start of day one of the 2wSR. The PICO question was selected for relevance to currently funded projects, anticipating that it would involve a moderately sized SR (2,000−3,000 search results and 10−20 included studies), although this was not verified in advance. The topic of the SR was purposefully well focused with a single, narrowly defined PICO question.

2.3. Outcomes

The outcomes included the following:

1) Time required to complete the 2wSR, from disclosure of PICO to complete draft of sufficiently high quality to circulate for feedback;
2) Time to complete individual SR tasks;
3) Facilitators and barriers to SR progress;
4) Peer reviewer feedback on the SR manuscript.

2.4. Data collection

For the first two outcomes, systematic reviewers recorded the calendar day(s) on which the task was started and completed and the time taken for each task in minutes; timing was paused if a break from the task exceeded 5 minutes. Data were recorded in Excel.

Facilitators and barriers were noted by the reviewers (J.C. and A.M.S.) as they arose and discussed during daily administrative meetings. Each meeting covered the following questions:

1. What tasks were completed the previous day?
2. What facilitators were identified, and how did they help?
3. What barriers were identified, and what (if anything) was helpful to overcome them?

Table 1. Characteristics of the completed SR [5]

Task	Evaluation measurement	Number
Systematic search	Number of databases searched	3
	Number of trial registries searched	2
Deduplicate	Number of records to be deduplicated	1694
	Number of records left after deduplication	1,381
Screening title and abstract	Number of studies to screen	1,381
Find full texts	Number of full texts required	48
Screening full texts	Number of full texts for screening	48
Data extraction	Number of full-text articles extracted (characteristics of studies, outcomes)	8
Assess risk of bias	Number of full-text articles requiring Risk of Bias assessments	8
Analysis	Number of full-text articles qualitatively synthesized	8
Analysis	Number of full-text articles meta-analyzed	7

4. What tasks were scheduled for today?

Two observers (A.B.B. and P.S.) sat in at the administrative meetings to take additional notes.

Peer reviewer feedback on the SR manuscript was provided by two journals. The number of comments received, sections of the manuscript to which they pertained, and the number of changes made in response were recorded in Excel. The time required to make changes in response to feedback was recorded in minutes.

2.5. Analysis

Data for the time to completion of the full SR and individual tasks were analyzed quantitatively. The total time (minutes) taken by individuals to complete each task was added, and aggregate time was reported in person-minutes and person-hours. Calendar days on which each task was completed were reported as recorded by the systematic reviewers.

Facilitators and barriers were mapped onto the Theoretical Domains Framework's [11] domains and constructs by discussion and consensus. One domain (Technology and Infrastructure) and one construct (Technical issues) were added to the Framework, to capture the focus of this project.

Peer reviewers' comments that did not request changes, and general comments (e.g., those summarizing the content of the manuscript) were removed, and the remaining comments were edited for brevity or clarity. The number of comments, the number of changes made in response, and the time required to make those changes were summed up.

3. Results

3.1. Time to complete the full systematic review

The SR was completed in 12 calendar days (between 21 January and 1 February 2019), working across a 5-day work week and a 4-day work week (because of a public holiday) (Table 4). A standard work week at our university is 37.5 hours (7.5 hours a day).

The SR commenced on 21 January 2019, with an all-author meeting (A.M.S., J.C., C.D.M., and P.G.) during which two authors (P.G. and C.D.M.) disclosed the PICO question to the rest of the team (A.M.S. and J.C.). The work on the protocol commenced subsequently, and searches for previous SRs on the topic were conducted, search strategy for the SR was designed and run, and data extraction forms were generated. During the first week of the 2wSR (21 January – 25 January), the search results were screened, full texts were obtained for most of the references, and risk of bias assessment commenced. The protocol was also completed—and its background and methods sections were transformed into an early draft of the SR manuscript. During the second week (29 January – 1 February), the remaining full texts were obtained, citation analysis was conducted, and its results were screened, risk of bias and data extractions were finalized, and the results were meta-analyzed. The results, discussion, and conclusion sections of the SR manuscript were written, and on 1 February, the draft of the manuscript was circulated to colleagues for feedback. On 5 February, the SR manuscript was submitted to a journal.

Four authors (A.M.S., J.C., C.D.M., and P.G.) required 61 hours to complete all SR tasks, from disclosure of the PICO (9:30am on Monday, 21 January) to the completion of a draft manuscript of sufficiently high quality to circulate for feedback (12:10pm on Friday, 1 February). The SR was therefore completed in 9 working days.

Formatting and revising the manuscript to the requirements of the first journal to which it was submitted required an additional 5 hours, for a total of 66 person-hours. The final, publishable version of the review required an additional 5 hours for a total of 71 person-hours.

3.2. Time to complete individual SR tasks

The time spent on each SR task is reported in Table 4. The most time consuming of all tasks were the daily administrative meetings (680 person-minutes). Among the SR tasks, the five most time-consuming tasks required

Table 2. List of SR tasks and tool(s) used

No.	Task	Description	Tool(s) used
0	Daily administrative meetings	Short daily meetings to review progress, discussion points were work done on previous day; work to be done this day; problems that may impede SR progress (barriers or facilitators for reporting in this study were also discussed)	—
1	Formulate review question	Decide on the research question of the review	—
2	Find previous or upcoming SR	Search for an SR that answers or in the future will answer the same question	—
3	Write the protocol	Design objective, reproducible, sound methods for the systematic review	Template
4	Design systematic search	Decide on databases and keywords to find all relevant trials	SRA—Word Frequency Analyzer, The Search Refiner
5	Design data extraction form and pilot	Design Excel forms for extracting study characteristics and test their usefulness/applicability	—
6	Run systematic search	Convert and run PubMed/MEDLINE search in all other databases	SRA—Polyglot Search Translator
7	Deduplicate results	Remove duplicate citations	SRA—Deduplicator, EndNote
8	Screen abstracts	Screen titles and abstracts, exclude irrelevant citations, resolve disputes	SRA Helper, RobotSearch
9	Obtain full text	Download, request copies from authors, interlibrary loan, etc.	Endnote; SRA Helper; SARA
10	Screen full text	Screen full text of articles, exclude irrelevant citations, resolve disputes	SRA Helper
11	Screen trial registries	Based on title and text in the trial registry entry: exclude irrelevant citations, dispute resolutions	—
12	Citation analysis	Follow citations, cited and citing, from included studies to find additional relevant studies	—
13	Screen citation analysis	Based on titles/abstracts and the full text of articles: exclude irrelevant citations, resolve disputes	SRA Helper
14	Extract data	Extract outcome numbers and associate with trial arm	RevMan
15	Assess risk of bias	Assess the potential biases in included trials	RobotReviewer
16	Synthesize and meta-analyze data	Convert extracted data to common representation (usually mean and SD), statistically combine the results (meta-analysis)	RevMan
17	Update systematic search	Repeat the search to find new literature published since the initial search.	See 6, 7, 8, 9
18	Write systematic review	Produce draft of the manuscript	SRA—RevMan Replicant
19	Revise manuscript for submission	Revise manuscript to meet journal requirements and standards	—

SRA, Systematic Review Automation; SARA, System for Automatically Requesting Articles.
Adapted from Tsafnat et al. 2014 [8].

approximately four or more person-hours each: data extraction (430 person minutes), write-up of the SR (417 person-minutes), title/abstract screen (369 person-minutes), risk of bias assessment (322 person-minutes), and full-text screening (227 person-minutes).

The least time-consuming tasks included obtaining full texts of articles (59 person-minutes), searches (58 person-minutes), citation analysis (30 person-minutes), data synthesis (30 person-minutes), and deduplication of search results (21 person-minutes).

Table 3. Automation tools used in the 2wSR

SRA tool	SR task	Description
SRA - Word Frequency Analyzer	Design systematic search	Accelerates designing a search by counting the number of times a word or phrase appears in a selected group of articles. Words that appear frequently should be used in the systematic search. http://sr-accelerator.com/#/help/wordfreq
The Search Refiner	Design systematic search	Accelerates designing a search by checking the recall (number of relevant studies found) and precision (number of irrelevant studies found) for each term in the search string and then displays it visually. Used to quickly determine which terms should be removed from the search string.
SRA - Polyglot Search Translator	Run systematic search	Accelerates running a search by converting a PubMed or Ovid Medline search to the correct syntax to be run in other databases. http://sr-accelerator.com/#/polyglot
SRA - De-duplicator	Deduplicate	Automates most of the deduplication process by identifying and removing the same study from a group of uploaded records. It is designed to be cautious so some duplicates will remain, which will require removing manually. http://sr-accelerator.com/#/help/dedupe
SRA Helper	Screen abstracts and obtain full texts	Accelerates screening and obtaining full texts by assigning to groups to be performed with a hot key. Hot keys are also assigned to search a list of prespecified locations to attempt to find the full texts of articles. http://sr-accelerator.com/#/sra-helper
RobotSearch	Screen abstracts	Automate citation screening by identifying the studies that are obviously not RCTs from a group of search results. Removes them, leaving a pool of potential RCTs to be screened [9]. https://robotsearch.vortext.systems/
Endnote	Screen abstracts, obtain full texts and write up SR	Accelerate multiple tasks and it assists with reference management. Useful for storing search results, finding full texts, sorting into groups during screening, and to insert references into the manuscript. https://endnote.com/ (N.B. proprietary software)
SARA	Obtain full texts	Automates requesting full-text articles to the library by requesting all needed full texts with a single request, which normally these requests need to be processed and sent one at a time. Tool (available within SRA): http://sr-accelerator.com/#/
RobotReviewer	Assess Risk of Bias	Accelerates assessing risk of bias on 4 of the 7 Risk of Bias domains by highlighting the supporting phrases in the PDF of the original paper. A check of the assessments is recommended, although the process is drastically speeded up [10]. https://robotreviewer.vortext.systems/
SRA — RevMan Replicant	Write-up SR	Accelerates the writing of a results section by having the computer write a first draft of the results section from the forest plots in a RevMan file. This draft can then be used as a start point to speed up the writing of the results. (Must be logged in access) http://sr-accelerator.com/#/replicant

RCT, randomized controlled trials; SR, systematic review; SRA, Systematic Review Automation.

The tasks of longest duration in calendar days included daily administrative meetings (held all 9 days), obtaining full texts and screening of full texts (5 days each), screening of abstracts (4 days; while most of the screening was performed on days 1 and 2, a PICO amendment required a rescreen on days 3 and 4), and writing of the manuscript (4 days). Four SR tasks were completed in one calendar day each: formulate PICO question, find previous or upcoming SRs, design systematic search, and deduplicate results.

3.3. Facilitators and barriers

Facilitators and barriers to completing the 2wSR mapped onto the following domains in the Theoretical Domains Framework: knowledge; skills; memory, attention, and decision process; environmental context and resources; and technology and infrastructure (a domain added to the original Framework to capture the focus of this project) (Table 5).

In the knowledge domain, prior knowledge of excluded interventions was an important facilitator as it increased the speed of literature screening. In the skills domain, the SR completion was facilitated by the extensive methodological expertise, adaption of existing data extraction sheets from prior projects, shortened timelines between screening decisions and dispute resolution thus minimizing 'mental reload' time, and the use of a validation set of articles. Although the reviewers found it challenging to *fully* block off time to focus on this project, focusing *predominantly* on the SR was attainable and facilitated its rapid completion. Daily administrative meetings allowed addressing issues as they arose, and writing the protocol in past tense accelerated its conversion into the manuscript. Environmental stressors such as noise

Table 4. Personnel and time required for each SR task

Task no.	Tasks	Number and authors (initials) involved	Total person time (mins and percent)	Date task started (2019)	Date task finished (2019)
0	Daily administrative meetings	4 Authors – (A.M.S., J.C., C.D.M., and P.G.)	685 (19%)	21 Jan	1 Feb
1	Formulate review question	4 Authors – (A.M.S., J.C., C.D.M., and P.G.)	120 (3%)	21 Jan	21 Jan
2	Find previous or upcoming SRs	1 Author – (J.C.)	63 (2%)	21 Jan	21 Jan
3	Write the protocol	4 Authors – (A.M.S., J.C., C.D.M., and P.G.)	175 (5%)	21 Jan	23 Jan
4	Design systematic search	1 Author – (J.C.)	109 (3%)	21 Jan	21 Jan
5	Design data extraction form and pilot	1 Author – (A.M.S.)	72 (2%)	21 Jan	22 Jan
6	Run systematic search	1 Author – (J.C.)	72 (2%)	21 Jan	24 Jan
7	Deduplicate results	1 Author – (J.C.)	16 (0%)	21 Jan	21 Jan
8	Screen abstracts	2 Authors – (A.M.S. and J.C.)	404 (11%)	21 Jan	24 Jan
9	Obtain full text	2 Authors (A.M.S. and J.C.)	41 (1%)	22 Jan	29 Jan
10	Screen full text	2 Authors (A.M.S. and J.C.)	187 (5%)	23 Jan	29 Jan
11	Screen trial registries	2 Authors (A.M.S. and J.C.)	123 (3%)	24 Jan	24 Jan
12	Citation analysis	1 Author (J.C.)	30 (1%)	24 Jan	24 Jan
13	Screen citation analysis	2 Authors – (A.M.S. and J.C.)	171 (5%)	25 Jan	29 Jan
14	Extract data.	2 Authors (A.M.S. and J.C.)	461 (13%)	25 Jan	30 Jan
15	Assess Risk of Bias	2 Authors (A.M.S. and J.C.)	323 (9%)	25 Jan	30 Jan
16	Synthesis and meta-analysis	2 Authors (A.M.S. and P.G.)	167 (5%)	31 Jan	1 Feb
17	Update systematic search	1 Author – (J.C.)	NA	NA	NA
18	Write-up of SR	4 Authors – (A.M.S., J.C., C.D.M., and P.G.)	428 (12%)	29 Jan	1 Feb
NA	Total time	4 Authors – (A.M.S., J.C., C.D.M., and P.G.)	3647		
Full SR draft of sufficient quality to circulate for feedback: 3647 person-minutes or 61 person-hours					
19	Revise and format manuscript for journal submission	4 Authors – (A.M.S., J.C., C.D.M., and P.G.)	325	5 Feb	5 Feb
Total time to submit to journal: 3972 person-minutes or 66 person-hours					
20	Meeting to discuss feedback from first set of peer reviewers + revisions	3 Authors – (A.M.S., J.C., and P.G.)	173	13 June	13 June
21	Meeting to discuss feedback from second set of peer reviewers + revisions	2 Authors – (A.M.S. and P.G.)	121	25 July	25 July
Total time to manuscript accepted for publication: 4266 person-minutes or 71 person-hours					

were mitigated by shutting office doors and noise-canceling headphones, and resources unavailable at our library were obtained from colleagues with access to other libraries. An important facilitator was the physical proximity of the systematic reviewers' offices to each other, enabling ongoing communication, as was the pre-existing familiarity with the existing automation tools and their uses.

The key barrier in the knowledge domain was the absence of clinical knowledge by the screeners, and in the skills domain, an omission of the check of agreement of screening decisions. In the memory, attention, and decision process domain, we found it difficult to fully block off time to focus solely on this project, as each of us was involved in other projects with competing priorities and deadlines. Environmental context and resource barriers involved noisy surroundings (the project took place during a teaching semester at our university, and construction was occurring in the building at the time), and resource unavailability (incomplete reporting by published studies, unavailability of full-texts through our library). Technology and infrastructure barriers centered around website glitches, software incompatibilities, and software limitations (e.g., automation of only part of the task or operationality only in the English language).

In the process of identifying facilitators and barriers to completing the SR, we also identified several SRA tools whose development could further accelerate the systematic review process. Some of these tools involve enhanced integration between tools that already exist, whereas others are stand-alone, new tools (Table 6).

Table 5. Facilitators and barriers to completing a 2wSR

Domain	Construct	Examples
Facilitators		
Clinical knowledge	Content knowledge	Prior knowledge of excluded interventions from the SR (e.g., antibiotics) sped up the screening for inclusion/exclusion
Skills	Methodological skills	Extensive methodological expertise in systematic reviews
		Reuse/amendment of data extraction forms designed for previous projects
		Shortened timelines minimize ''mental reload'' time (e.g., completing screening and dispute resolution on the same day, facilitates with recall of the screening decisions, obviating the need to check reasoning for exclusions/inclusions)
		Validation set of articles was used to generate the search strategy
Memory, attention, and decision processes	Decision-making (planning)	Blocked off time from other projects to focus on the SR project
		Daily administrative meetings to address issues as they arise
		Writing protocol in past tense (not future tense) to facilitate conversion of introduction and methods to manuscript with minimal amendments
Environmental context and resources	Environmental (de)stressors	Working in offices (rather than open-plan environment), shutting the office door, and use of noise-canceling headphones
	Resource availability	Close physical location of systematic reviewers' offices, allowing ongoing communication and flexibility to resolve issues in real time
		Access to libraries in addition to our institutional library
		Quick response time to queries by authors of included studies
Technology and infrastructure	Technical issues	Pre-existing knowledge and ability to use the automation tools (e.g., SRA-Helper, Word Frequency Analyzer, etc.) eliminated 'learning curve' and time
		Use of google translate to check articles in foreign languages for inclusion
Barriers		
Clinical knowledge	Content knowledge	Lack of clinical expertise by systematic reviewers who screened the literature
Skills	Methodological skills	Omission of the standard check of agreement in title/abstract screening decisions for the first 50 title/abstracts extended the dispute resolution time
Memory, attention, and decision processes	Decision-making (planning)	Difficulty blocking off time to work only on a single project, as deadlines or 'emergencies' for other projects arose and required attention and time
Environmental context and resources	Environmental stressors	Noisy surroundings (teaching semester, construction in the building)
	Resource unavailability	Incomplete reporting by published studies, necessitating author contact
		Full texts of some studies were unavailable from our university library
Technology and infrastructure	Technical issues	Websites nonoperational (e.g., WHO ICTRP and Cochrane library were down)
		Software incompatibilities (e.g., Robot Reviewer and Internet Explorer)
		Slow internet (e.g., websites loading very slowly)
		Software automates only parts of a task (e.g., RobotReviewer only assesses 4 of 7 risk of bias domains; Polyglot does not translate MeSH terms)
		Software only operational in English language (e.g., RobotReviewer)
		Poorly phrased output produced by automation tools (RevMan Replicant automated text)
		SRA Helper (Endnote Helper) does not permit ''highlighting'' of included/excluded terms (for example, as in Covidence)

3.4. Peer reviewer feedback on the SR manuscript

The manuscript was submitted to four journals and rejected without peer review by two journals and by a third journal with peer review feedback, and accepted by a fourth journal. Timelines for decisions ranged from 1 week to 9 weeks (Table A1).

The first set of peer reviewer feedback included comments from two peer reviewers (Table A2). Thirty-four comments requesting changes were received; each section of the manuscript received comments, with the majority (11/34) focusing on the methods section. Three SR authors

Table 6. SRA tools for future development

SR task	Issue or gap identified	Potential SRA tool
Design systematic search	Collating common terms from relevant articles is time consuming	Search term collator – automatically extracts common terms for a key set of articles and allows grouping of similar ones
Run systematic search	Modifying index terms for individual databases is time consuming and error prone	Index term converter – automatically converts index terms between databases (e.g., MeSH terms to Emtree terms)
Screen abstracts	Learning time can be time consuming for noncontent experts	Topic thesaurus – highlights prespecified terms in the title and abstract to aid in exclusion or inclusion decisions
Citation analysis	Time consuming to identify and collate all citing and cited articles of included studies	Citation collator - automatically identifies and collates all articles in a reference list or those that cite the included studies
Data Extraction	Authors of included studies often need to be contacted for additional information	Author contactor - automatically populates from the meta-data of included papers a tool that will email authors automatically and track responses
Data Extraction	Lack of integration between various software	SRA tool integrator – automatically exchanges data between various software currently used in SRs
Update systematic search	Removing articles already screened (in the initial search) is a manual and time-consuming process	Citation tracker – keeps track of all articles sent to screening teams and automatically removes any updated search results that have been previously screened
Write SR	Table of included studies are time consuming to create	A tool to automatically generate a table of included studies from Excel or RevMan that allows for reformatting depending on journal requirements
Write SR	Time consuming to generate text about PRISMA flowchart for publication	PRSIMA Flowchart Text Generator – a tool to automatically generate accompanying text from the PRISMA flowchart in RevMan

SR, systematic review; SRA, Systematic Review Automation.

(P.G., J.C., and A.M.S.) met to discuss how best to address the suggested changes: 17 changes were made fully or partially. The meeting and the revisions required 173 person-minutes (Table 4)

The second set of peer reviewer feedback comprised comments from two peer reviewers (Table A3), requesting 13 changes; the results section received most feedback (4/13 comments). The introduction section and tables received no comments, whereas the remaining sections received from one to four comments each. A meeting to discuss the revisions and the revisions themselves (8 changes) required 121 person-minutes.

4. Discussion

Our SR team included four members (compared to a median of five for an SR), required screening of 1,694 studies (slightly less than the median of 1,781), and was completed in 61 person-hours or 9 working days. The time to journal submission was 66 person-hours (16 calendar days), which represents a considerable improvement on the median time to journal submission of 41 weeks (Borah et al. 2017). The final, publishable version of the manuscript required 71 person-hours.

A recent study, evaluating the time logs of 12 simulated SRs found that the average time to SR completion was 463 days (66 weeks) and 881 person-hours [12]. The study reported the time consumed by each task: selecting studies 26% (229 hours) of the total person-hours per SR; collecting data 24% (211 hours); preparing report 23% (202 hours); conducting meta-analysis 17% (149 hours); and descriptive synthesis 6% (52 hours). In comparison, in the present study, selecting studies consumed 16% of the time (552 minutes or 9 hours), extracting data 13% (461 minutes or 8 hours), and conducting meta-analysis 5% (167 minutes or 3 hours), and no descriptive syntheses were conducted. This suggests that systematic reviewers focused on a single SR, who prioritize the SR over other projects, and communicate in real-time may considerably reduce time to completion both for the individual SR tasks and the entire SR. An additional benefit of the 2wSR approach over conventional SR methods is the efficiency gain realized from not having to rerun out-of-date searches and subsequently incorporating additional studies at the completion of the SR.

Our SR team included experienced systematic reviewers with complementary areas of specialization, used SRA tools where possible, blocked off time from other projects for the duration of the SR; a daily meeting was also held to identify the barriers and facilitators to completing the SR. These elements can be adopted and replicated by other SR teams, although the generally low adoption of SRA

tools [3,4] suggests that this element may be most challenging to replicate by some teams.

However, the importance of acquiring knowledge of the existing SRA tools and facility with their use—as well as knowledge of their limitations—cannot be overstated. This is because, at this point in time, very few of the existing SR tools are sufficiently developed to completely replace a user—most tools can only assist a user in completing the task (one of the few exceptions to this is the RobotSearch tool). Nevertheless, incorporating the SRA tools into the SR workflow allowed us to enhance our speed and work in a more targeted way. The tools not only sped up the process but also removed some of the 'tedium' from the more repetitive tasks, allowing the users to retain their focus. Indeed, we found that generally, the least time-consuming SR tasks were those for which automation tools were available, and conversely, the most time-consuming tasks were those for which automation tools were either unavailable or available to automate only part of the task. As automation tools are becoming extensively tested—including the Deduplicator [13] and the Polyglot Search Translator [14] used in our review—and the increase in efficiency associated with their use becomes more apparent, the willingness to adopt them may increase.

The moderate size of the SR, intervention question (rather than e.g., diagnostic or prognosis), well-focused scope and narrowly defined PICO question, and inclusion of only RCTs contributed to its completion in 2 weeks. Nevertheless, it is possible to realize considerable time savings over the 41-week median to SR publication, even with larger and more complex SRs.

A second 2wSR was conducted approximately 5 months after the first, of the RCTs evaluating the impact of self-management interventions, in men with lower urinary tract symptoms. The same processes were followed, and the team consisted of four experienced systematic reviewers (two from the first 2wSR (P.G. and J.C.), and two other researchers from our institute). The reviewers were experienced, familiar with SRA tools, and held daily meetings to address issues. Barriers and facilitators were recorded (Table A4). The systematic search found 2872 references, which were screened, 38 articles reporting 25 studies were included. The SR was completed (i.e., a draft manuscript of sufficient quality to circulate for feedback) in 11 working days.

In general, both 2wSRs encountered similar facilitators and barriers. Both teams found the short daily meetings, close physical proximity between team members' offices, and short time lapse between tasks such as screening and dispute resolution to be very helpful. In both cases, the reviewers found it challenging to fully block off 2 weeks from other projects, blocking off a half-day every second day during a 2wSR, to attend to other projects may be helpful. Finally, both teams also disliked some of the output produced by automation software, but both also agreed that it was easier and less error prone to edit automatically produced text than to write it from the start.

However, although the advantages to using the 2wSR approach are evident, disadvantages may be less so. The 2wSR approach requires staff members to focus almost entirely on the SR. This means that they are not available to contribute to other projects they are involved in—in our case, that meant, for example, that contribution to another manuscript that required to be revised and resubmitted for publication was delayed. Teams adopting the 2wSR approach need to be aware that this approach may impact on progress of other projects they are involved in. Future work in this area will include trialling methods to address this issue and targeting both larger and more complex SRs, which will require larger teams. We will monitor whether the processes described here can be adapted to such work. As further SRA tools are developed—for example those identified in Table 6—we will also integrate them into our processes and assess their impact on our workflows.

4.1. Strengths and limitations

This manuscript describes the results of the first case study in conducting a 2wSR. The greatest strength of the study is its novelty—to our knowledge, this is the first study of its kind. Moreover, the study contributes to the body of knowledge on how long individual SR tasks take and is one of the few to integrate multiple SRA tools into the SR, rather than focusing on the impact of integrating a single one. However, because the results are based on a single SR project, they may not be generalizable to other SR projects or to other teams conducting SRs. Moreover, the SR had a clearly defined focus with a narrow PICO question, which may not be representative of other SRs.

Nevertheless, we have repeated the 2wSR process with a larger SR (8 included studies in the initial 2wSR and 25 including studies in the second), finding similar barriers and facilitators in both cases. We are therefore confident that the process is adoptable and adaptable to other SRs and offers a potential to realize considerable time and efficiency savings.

5. Conclusion

A small and experienced SR team, using SRA tools who have protected time to focus solely on the SR can complete a moderately sized SR in 2 weeks.

Acknowledgments

The authors would like to thank Kathryn Kaiser for providing additional data from the Borah et al. 2017 study and Ba' Pham for feedback on the manuscript.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.01.008.

References

[1] Borah R, Brown AW, Capers PL, Kaiser KA. Analysis of the time and workers needed to conduct systematic reviews of medical interventions using data from the PROSPERO registry. BMJ Open 2017;7(2):e012545.

[2] Beller E, Clark J, Tsafnat G, Adams C, Diehl H, Lund H, et al. Making progress with the automation of systematic reviews: principles of the international collaboration for the automation of systematic reviews (ICASR). Syst Rev 2018;7(1):77.

[3] O'Connor AM, Tsafnat G, Gilbert SB, Thayer KA, Wolfe MS. Moving toward the automation of the systematic review process: a summary of discussions at the second meeting of International Collaboration for the Automation of Systematic Reviews (ICASR). Syst Rev 2018;7(1):3.

[4] van Altena AJ, Spijker R, Olabarriaga SD. Usage of automation tools in systematic reviews. Res Synth Methods 2019;10(1):72−82.

[5] Scott A, Clark J, Del Mar C, Glasziou P. Increased fluid intake to prevent urinary tract infections: systematic review and meta-analysis. Br J Gen Pract 2020. https://doi.org/10.3399/bjgp20X708125.

[6] Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Chichester, West Sussex; Hoboken NJ: Wiley-Blackwell; 2011.

[7] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535.

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

[9] Marshall IJ, Noel-Storr A, Kuiper J, Thomas J, Wallace BC. Machine learning for identifying Randomized Controlled Trials: an evaluation and practitioner's guide. Res Synth Methods 2018;9(4):602−14.

[10] Marshall IJ, Kuiper J, Wallace BC. RobotReviewer: evaluation of a system for automatically assessing bias in clinical trials. J Am Med Inform Assoc 2016;23(1):193−201.

[11] Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005;14:26−33.

[12] 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.

[13] Rathbone J, Carter M, Hoffmann T, Glasziou P. Better duplicate detection for systematic reviewers: evaluation of Systematic Review Assistant-Deduplication Module. Syst Rev 2015;4:6.

[14] Clark J, Sanders S, Carter M, Honeyman D, Cleo G, Auld Y, et al. Improving the translation of search strategies using the Polyglot Search Translator: a randomised controlled trial. J Am Libr Assoc 2019. In press.