REVIEW

The role of open research in improving the standards of evidence synthesis: current challenges and potential solutions in systematic reviews [version 1; peer review: awaiting peer review]

Eirini Martinou¹, Angeliki Angelidi²

¹Faculty of Health and Medical Sciences, University of Surrey, Guildford, Surrey, GU2 7XH, UK
²Broad Institute of MIT and Harvard, Harvard Medical School, Boston, MA 02142, USA

Abstract

Systematic reviews (SRs) and meta-analyses (MAs) are the cornerstone of evidence-based medicine and are placed at the top of the level-of-evidence pyramid. To date, there are several methodological resources available from international organizations such as the Cochrane Collaboration that aim to aid researchers in conducting high-quality secondary research and promoting reproducibility, transparency and scientific rigour. Nevertheless, researchers still face challenges in most stages of evidence synthesis. Open research and the FAIR (findability, accessibility, interoperability, and reusability) principles are rising initiatives being increasingly implemented in primary research. However, their beneficial role in secondary research is less emphasized. This article addresses how the challenges commonly faced during evidence synthesis research could be overcome using open research practices and currently available open research tools. Despite the phenomenally simple SR workflow, researchers still find tasks such as framing the SR research question, search strategy development, data extraction, and assessing for bias, challenging. The implementation of FAIR practices, including prospective registration at the PROSPERO database, abiding with the PRISMA guidelines, and making all SR data openly available could have significant benefits in avoiding duplication of effort and reducing research waste while improving the reporting standards of SRs. Additionally, this article highlights the need for further education in open research culture to overcome ethical and motivational barriers in implementing open research practices in evidence synthesis. Finally, in the era of technological breakthroughs, artificial intelligence may eventually be incorporated into the process of SRs and should abide by the FAIR standards for open research.
Keywords
evidence synthesis, open research, FAIR principles, systematic review

Corresponding authors: Eirini Martinou (em01023@surrey.ac.uk), Angeliki Angelidi (aangelid@broadinstitute.org)

Author roles: Martinou E: Conceptualization, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing;
Angelidi A: Conceptualization, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2022 Martinou E and Angelidi A. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Martinou E and Angelidi A. The role of open research in improving the standards of evidence synthesis: current challenges and potential solutions in systematic reviews [version 1; peer review: awaiting peer review] F1000Research 2022, 11:1435 https://doi.org/10.12688/f1000research.127179.1

First published: 05 Dec 2022, 11:1435 https://doi.org/10.12688/f1000research.127179.1
Introduction
Evidence synthesis refers to any method that identifies, selects, and combines results from multiple studies. It includes reviews (narrative, systematic, rapid, scoping, umbrella) and meta-analyses (MAs). Narrative reviews are used widely in evidence synthesis; however, they tend to be descriptive and do not have a standardized methodology or established protocol. On the contrary, a systematic review (SR), according to the Cochrane Collaboration, is defined as a review that identifies, appraises, and synthesizes all the available evidence to answer a specific research question under pre-specified eligibility criteria. SRs aim to summarize the available evidence comprehensively and transparently while minimizing bias and enhancing the reliability of the related conclusions.

Since the first publication of a SR by James Lind in 1753, there has been tremendous growth in this field with a more than 20-fold increase in the number of SRs indexed over the last two decades. An observational study by Hoffman et al. estimated that approximately 80 SRs were being published per day in the last 20 years and over 200,000 might now be available. Within the field of clinical practice, SRs and MAs are placed at the top of the level-of-evidence pyramid as they form the basis of clinical practice guidelines (Figure 1). SRs are considered the cornerstone of evidence-based medicine; however, their role is expanded in other scientific areas including social sciences and basic science research as they promote the current understanding and facilitate the identification of research gaps.

The use of a precise and explicit methodology is of paramount importance in conducting SRs and therefore, several methodological resources have been developed to aid researchers in this field. To date, several organizations such as the National Institute of Health and Clinical Excellence (NICE) in the UK, the Evidence-based Practice Centre Program (EPC) in the Netherlands, and the Centre for Evidence-Based Medicine (CEBM) in Oxford, UK, have developed comprehensive methodologies and guidelines for conducting high-quality SRs. These resources provide researchers with a systematic approach to reviewing, evaluating, and synthesizing the available evidence, ensuring that their conclusions are based on robust and transparent evidence.
the USA, and the international Campbell and Cochrane Collaborations have been established and are dedicated to developing tools and methods guides for research synthesis and evidence reports (such as SR). However, despite the widely available resources, researchers still face challenges in transparency when conducting SRs.

Open research (also referred to as “open science”) is a term that encompasses the whole research process making the whole research process transparent. The principles of findability, accessibility, interoperability and reusability (FAIR) were formulated in 2014 and were firstly designed only for data-sharing practices. However, FAIR principles have been used more broadly in research policies and have become the cornerstone of research data management and stewardship. FAIR principles apply not only to data but also to research workflows and protocols used in research studies. Moreover, the application of FAIR principles could be expanded beyond primary research studies to evidence synthesis research. Our article addresses open research and SRs using the FAIR principles as a basis. Open research is a rapidly evolving shift in research culture to promote transparency and openness in every field of research and is supported by several organizations such as UK Research and Innovation (UKRI), the National Institute of Health Research (NIHR), and the United Nations Educational, Scientific and Cultural Organization (UNESCO). Interestingly, although open research practice is highly promoted, the use of FAIR principles is still lacking in systematic reviews making them challenging to be conducted and their quality questionable.

This article aims to highlight the current challenges and barriers encountered while conducting SR and suggest potential solutions by emphasizing the importance of open research practices. Moreover, this manuscript describes the role and future perspectives of artificial intelligence in the FAIRness of evidence synthesis. In our article, we first describe the main steps in the SR pipeline. We further review the challenges the researchers may face during these steps of evidence synthesis research and suggest potential solutions by using open research tools and abiding by open research principles. Secondly, we summarize the difficulties and barriers that may impede the full implementation of FAIR principles in SRs, and we propose potential solutions to overcome these challenges. Lastly, we discuss future perspectives regarding the role of using automation tools in SRs within the context of FAIR principles.

Overview of systematic review workflow

SRs as well as MAs are characterized by an explicit and robust methodology which consists of several consecutive steps starting from framing the research question followed by identifying relevant work, appraising the quality of included studies, summarizing the evidence, and interpreting the results. However, researchers while conducting an SR, may experience many challenges despite their phenomenally simple workflow. FAIR principles refer to the fact that research data should display the following characteristics: data should be registered in a searchable source and assigned a unique identifier, be retrievable ideally through automated procedures, be available to be combined to maximize their value, and lastly be described in detail to facilitate their reuse (Figure 2). The movement toward an open research culture that supports the FAIR principles in data sharing may have important implications in the conduct, methodology and reporting of SRs and other evidence synthesis studies. Below we discuss the potential obstacles that researchers may often encounter when conducting SRs and suggest potential solutions.

Figure 2. Findability, Accessibility, Interoperability and Reusability in evidence synthesis.
encounter while conducting an SR and how they could overcome these barriers by using open research practices and open relevant tools

### Challenges in evidence synthesis and open research solutions

#### Framing the research question

Conceptualizing an SR and formulating the research question is the most important step in evidence synthesis studies. A significant challenge that researchers face at this stage is to identify information regarding ongoing SRs addressing the same research question. Ioannides et al. have painted a picture in evidence synthesis research activity where there has been an explosion of SRs and MAs that are often redundant. Katsura et al. also suggest that multiple SRs on the same topic represent wasted efforts and they can be confusing and potentially misleading.

Interestingly, until 2011, SRs or meta-analyses were mostly available through search engines such as MEDLINE and EMBASE only after they have been completed and published. Researchers and healthcare evidence users were facing challenges in finding ongoing SRs as only a limited number of organizations including Cochrane Collaboration contained information related to SRs that were in progress. To overcome these challenges and promote FAIRness in evidence synthesis, the International Prospective Register of Systematic Reviews (PROSPERO) was launched in February 2011 as an open-access database where all health-related SRs are eligible for prospective registration. Additionally, the status of SRs is regularly updated and therefore researchers and healthcare professionals are able to find SRs which are in progress or are completed but not yet published. Open-access findability of ongoing SRs provides researchers and commissioners with an important tool to avoid duplication and reduces research waste. It is worth highlighting that the ability to promote FAIRness of ongoing SRs through PROSPERO does not only benefit researchers but also is important for clinicians and other healthcare professionals to remain up to date with related ongoing developments in healthcare.

Despite, the benefits of findability through the PROSPERO open access database and the increase in its use over the years, still, only a fraction of SRs were noted to be registered. Tawfik et al. by conducting a global survey of 270 authors, reported that only 50% of SRs were officially registered. Interestingly, the most common cause was a lack of awareness of the importance of open-access prospective registration of SRs, indicating the need for academic education in open research culture. To overcome this observation and encourage PROSPERO registration, the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement updated its criteria in 2020 for reporting SRs by moving the registration information to a separate checklist section to encourage open research practice within the global research community and promote FAIRness in SRs.

#### Identifying relevant work

To capture as many relevant citations as possible during an SR, a diligent and thorough search should be conducted. An extensive literature search is achieved using multiple electronic databases to identify primary studies according to a priori specified inclusion and exclusion criteria. Identifying relevant citations is a fundamental step for a reliable SR and includes the development of a meticulous Boolean strategy designed separately for each database. Subsequently, screening steps are performed based on the abstract and full text. A survey by Major et al. showed that the design of a search strategy and the performance of an extended literature search were reported to be the most challenging aspects that researchers encountered during an SR.

The development of a thorough Boolean strategy algorithm including all the possible vocabulary and synonyms is a relatively complex and time-consuming process which could be overcome by implementing open research principles in SRs through the current available open research tools (PROSPERO). A Boolean algorithm which is easily findable and accessible could be reused or modified by several other research teams who wish to conduct an SR on a different subject in the same research field to avoid duplication of research efforts. Interestingly, despite the presence of reporting guidelines such as PRISMA and the need for FAIRness and transparency in reporting the methodology of SRs, several studies have identified that in most SRs, a full Boolean search strategy was missing from the published articles. To further promote open research practices in SRs, one of the alterations noted in the updated PRISMA 2020 guidelines is the modification of the search term which now recommends that full search strategies for all databases should be reported. Additionally, the process of registration in the PROSPERO database requires the submission of the full search strategy; however, making this information publicly available prior to the publication is still optional. Recently, preliminary results from the REPRISE study by Page et al. showed that in a random SR sample less than 1% of the included studies reported an analytic search code which was used to generate results, highlighting that despite available open research tools and the revised PRISMA 2020 guidelines, further interventions are needed. The development of common rules
across the databases regarding the Boolean strategy design or the presence of easily accessible instructions could potentially aid researchers in tackling the challenging task of analytic vocabulary code design.

The screening phase of an SR aims to select potentially eligible studies for further assessment based on the prespecified inclusion and exclusion criteria. According to Cochrane guidance, two or more reviewers should independently perform this step, first by screening the title and abstract followed by a full-text assessment of the studies initially included. A common challenge that researchers face is the lack of consistent full-text availability leading them to proceed with individual communication with authors requesting full texts of the candidate manuscripts with variable responses. Studies have shown that data requests are ignored in up to 41% of cases. A potential solution to this challenge is proposed to be open-access publishing which is steadily increasing in popularity as articles that are freely accessible are more shared and ultimately benefit the distribution of knowledge. However, open-access publishing is associated with increasing article processing charges that place an increased financial burden, especially on unfunded and early-career researchers. Although journals and publishers should continue to encourage open-access publishing, we encourage them to consider alternate models of support to reduce the financial barriers that lesser-funded or unfunded researchers face. Furthermore, given that institutional affiliation contact details frequently change in a researcher’s career, the availability of research websites such as ResearchGate or other social networking sites for researchers could facilitate communication across the research community and promote accessibility to full text or data of a study. In addition, we believe researchers should be entitled to free access to articles related to the SR they are conducting to promote knowledge and improve the scientific rigour of the study.

*Data extraction and quality assessment of included studies*

This step encompasses data collection from the included full texts and is accompanied by a quality assessment of the included studies. These steps not only require a considerable amount of time and effort but also generate a vast amount of valuable secondary data which could be used and built upon to generate further new knowledge.

Data extraction constitutes a significant portion of an SR. Making all data extracted during SRs openly available could reduce unnecessary duplication of effort from other researchers and promote a standardized format in which data are presented in SRs. To increase SR data openness, the Systematic Review Data Repository (SRDR) was developed and launched in 2012 by the Brown University Evidence-based Practice Center and funded by Agency for Healthcare Quality and Research (AHRQ). This open-access resource aims to serve as an archive for SRs as well as an open-access data extraction tool. Saldanha et al. reported that since 2012, SRDR has gathered over 150 SRs and includes publicly available data from more than 15,000 studies covering a variety of health-related research subjects. Cochrane collaboration has also made initiatives to support accessibility in SR data and is committed to being fully open access by 2025.

Quality assessment of included studies and risk of bias reporting are key features of a high-quality quantitative SR and one of the most challenging stages in evidence synthesis. Risk of bias assessment requires a thorough judgement of the included studies which may introduce an element of inter-rater variability meaning that the same data may be judged differently between research groups. For this challenge, there are several available standardized open tools such as the Cochrane tools (RoB1, RoB2, ROBINS-I), especially in health-related themed SRs which aid researchers in this task. Interestingly, although the risk of bias is considered essential according to the PRISMA guidelines, several studies assessing the reporting quality of published SRs showed that risk of bias reporting is performed in less than 50%, raising concerns regarding their quality. A potential solution to this issue may be provided with the a priori registration of an SR in the PROSPERO open-access database which also requires a detailed study design, search strategy and protocol submission with the aim to minimize bias and promote high methodological standards and reproducibility. Studies by Sideri et al. and Allers et al. reported that SRs which were officially registered in the PROSPERO database appeared to demonstrate higher standards in their methodology and reporting quality. However, this improvement may not only be attributed to the early accessibility of the registered SRs’ design and protocols but also to the fact that authors may have invested more research time in a high-standard SR research design which may indirectly result in a more robust methodology.

Providing open access to all aspects of an SR including the risk of bias availability and raw extracted data could provide several benefits including a reduction in duplication of effort, improvement in the quality and efficiency of SRs as well as support for secondary analyses addressing additional research questions.

**Barriers to implementing FAIR principles**

Despite the undoubtedly significant numerous benefits of SR data sharing, several challenges exist in fully implementing FAIR principles related to motivational and ethical barriers. Ethical concerns, including ownership of data, intellectual
property of ideas, and recognition of data producers, are some of the barriers related to moral principles that may challenge data sharing and the widespread practice of FAIRness in research studies, including SRs. Zuiderwijk et al. performed an SR to investigate what inhibits researchers from openly sharing their data. Their study identified trust as a powerful and impactful driver for FAIR data sharing especially in a pre-publication state. Notably, Zuiderwijk et al. characterised fear over data control or recognition of researchers’ efforts on data generation as major trust inhibitors for open data sharing. This barrier is widely recognized, and several organizations have highlighted the need for appropriate systems for recognition so that acknowledgements are made as needed. For instance, Cochrane collaboration states that authors should “respect and acknowledge the source of the data”. Additionally, the SRDR has adopted data citation with digital object identifiers (DOIs). Another practice is the registered reports publication process which is being implemented by an increasing number of journals and publishers where study proposals and protocols are peer-reviewed and pre-accepted before the actual research is undertaken and results are produced. Registered reports are an increasingly adopted publication practice as it envisages promoting transparency, reproducibility and openness in data sharing in line with the transparency and openness promotion guidelines.

Personal drivers and intrinsic motivations have been identified by Zuiderwijk et al. as important factors impacting openness in data sharing. Scholars exhibit different character or belief-related traits in terms of being supportive or reserved in open data sharing. In an effort towards behavioural and academic culture change, the role of education in an open research culture is of paramount importance. Ioannidis et al. highlighted the need for a strong educational curriculum to equip researchers and meta-researchers with the knowledge of best scientific practices to promote FAIRness in meta-research. Furthermore, Universities and other academic institutions may consider rewarding meta-researchers who make their analytic code and research data publicly available as proposed by the 6th World Conference on Research Integrity and in line with the Hong Kong principles for assessing researchers to foster trustworthiness, rigour, and transparency.

Along with the various organizations which promote open research in evidence synthesis and the currently available open research tools, it is worth mentioning the key role that journal publishers and research funders play in enhancing data sharing and its reuse. In 2016, the International Committee of Medical Journal Editors (ICMJE) proposed that data from published clinical investigative trials need to be shared to promote transparency and reusability. However, to date, no relevant policy exists for evidence synthesis research. Modification of the interface of the existing SR registries to include detailed protocols, search codes and raw data may further promote the reusability of SR/MA-generated data and improve the standards of evidence synthesis reporting. Additionally, an increasing number of publications, such as F1000 and Systematic Reviews, provide clearer requirements for datasets to be shared in accordance with the FAIR principles.

Future directions: artificial intelligence and FAIRness in SRs

Undoubtedly, evidence synthesis research is growing exponentially, therefore the current robust but slow and resource-intensive process of the standard SRs is no longer sustainable. Therefore, several methods and tools have been developed through the use of artificial intelligence (AI) in order to aid the automation of SRs. AI includes machine learning (ML) which utilizes computer algorithms similar to logistic regression and natural language processing (NLP) which analyses a vast number of texts and extracts information. Both ML and NPL are commonly implemented technologies used in the semi-automated conduction of SRs. ML uses statistical predictive methods to calculate the likelihood that an article is relevant and is commonly used during the screening process whereas NPL analyses the semantic meaning and is used during the data extraction step of an SR.

Several automation tools have been developed that aim to execute the time-consuming and labour-intensive tasks of an SR such as search (e.g. Meta), article screening (e.g. Abstrackr), data extraction (e.g. ExaCT) and even automatic generation of PRISMA diagrams (e.g. PRISMA flow diagram generator). The significant advance in AI software as well as the desire for openness and FAIRness in meta-research has led to the initiation of the International Collaboration for the Automation of SRs (ICASR). In 2015, in the first ICASR meeting, a set of principles were established known as the Vienna principles which recommend the need for a collaborative multidisciplinary effort towards the automation of the evidence synthesis research process while highlighting that every automation technique should abide by the FAIR standards for open research and be freely available.

Despite the development of several automated platforms for SRs, the complete automation of the evidence synthesis process is not feasible at present, as no current AI tool can replace human judgement. Additionally, the benefits of the implementation of AI methods in SRs remain unclear and AI platforms still need to undergo validation and refinement. Nevertheless, AI tools are increasingly becoming incorporated into some of the steps of evidence synthesis until a fully automated process might be possible in the near future, (Figure 3). The conduction of evidence synthesis using AI in line with the FAIR principles and the utilization of computer science and health informatics may dramatically alter evidence synthesis research and eventually evidence-based medicine.
Conclusion

In conclusion, this review has highlighted the need for a transparent approach in evidence synthesis research in which implementing FAIR principles play a crucial role in tackling methodological challenges as well as improving the reporting standards of SRs. These may have broader benefits related to public health issues as data extracted from high-quality SRs could accelerate the conduction of rapid evidence-based reviews to produce information promptly, promote research dissemination and engage effective strategies, especially in global health crises such as major outbreaks or pandemics.\textsuperscript{37,62} Since existing SRs are often used by rapid reviews, making SR data as openly available as possible could provide vital answers at high standards and in a timely fashion for governments and policy makers in case of public health threats.\textsuperscript{63}

Finally, educating evidence synthesis researchers to follow the FAIR guidelines is of paramount importance for an academic cultural shift towards open research. AI may play an important role in the future of SRs and may aid in a more standardized and open evidence synthesis practice.

Data Statement

No data are associated with this article

Author contributions

Conceptualisation: EM, AA; Project administration: EM, AA; Writing of original draft: EM, AA; Visualisation: EM, AA; Review and Editing: EM, AA

Acknowledgements

The authors would like to thank Dr Daniele Kurtin and Professor Emily Farran, University of Surrey for their contribution during the editing process of the article (written consent was obtained)


46. Sideri S, Papageorgiou SN, Elades T.: Registration in the international prospective register of systematic reviews (PROSPERO) of systematic review protocols was associated with increased review quality. *J. Clin. Epidemiol.* 2018 Aug; 100: 103–110. Publisher Full Text


The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com