Intelligent use of artificial intelligence for systematic reviews of medical devices

Kelly Goodwin Burri
Stryker GmbH, Selzach, Switzerland

Correspondence to:
Kelly Goodwin Burri
Stryker GmbH
Dr Homer Stryker Strasse 1
2545 Selzach, Switzerland
kelly.goodwinburri@stryker.com

Abstract
Systematic literature reviews are an essential component of the medical device clinical evaluation process. The EU Medical Device Regulation requirement for regularly updated systematic literature searches will increase the burden on the medical writer to maintain and update systematic reviews for many systems and devices. Specialists in systematic reviews are beginning to adopt artificial intelligence tools that aim to optimise searches and streamline the review process. As these tools mature, the medical device writer tasked with a systematic review may want to consider the potential benefits of integrating them into their established systematic review process.

Systematic reviews are the foundation for evidence-based medicine and clinical guidelines. They are also an essential component of the clinical evaluation of medical devices marketed in Europe. The EU Medical Device Regulation 2017/745 (MDR) prescribes a systematic scientific literature review as part of the clinical evaluation process to "identify available clinical data relevant to the device and its intended purpose and any gaps in clinical evidence".[1] The EU MDR also includes regular screening of the scientific literature as part of the general methods and procedures for post-market clinical follow-up of marketed medical devices.

The systematic review process is time and resource intensive, requiring highly skilled reviewers to complete a series of very specialised manual and repetitive tasks. Add to that the reality of a vast and continuously expanding body of medical literature on which a review is based and, in some cases, the entire process takes so long to complete that a review may already be outdated by the time it is published. Various groups, including Cochrane, the recognised expert source on systematic reviews, acknowledge that it is not possible to keep all systematic reviews up-to-date and have developed guidance on when an update is appropriate.[2,3] However, for the medical device industry, the EU MDR dictates the frequency of these updates (e.g., annually for the highest risk Class III devices) as part of the ongoing clinical evaluation process. Device manufacturers understandably should have an interest in the development and implementation of new technologies to make the systematic review task faster and more efficient.

Artificial intelligence (AI) experts have zrealised the inherent challenges of the conventional systematic review process and are championing AI technology as the key to managing the flood of scientific literature.[4] AI has become prominent in the healthcare field, and there is now an emerging AI subspecialty specifically focussed on how to improve systematic reviews. Machine learning and natural language processing are current applications of AI that hold promise for evidence-based medicine to generate, update, and maintain an up-to-date synthesis of clinical data in a given field. So how exactly can AI improve the systematic review process? What kind of AI tools should the medical device writer be aware of? And should we expect machine learning to eventually relieve us of this clinical evaluation task altogether?

The conventional approach to systematic reviews

Systematic literature reviews can be broken down into several discrete tasks:[5]

1. Protocol – definition of review question, search query, and selection criteria
2. Search – conduct searches in relevant databases
3. Screen – initial publication selection based on title and abstract review; final selection based on full-text articles
4. Extract – extraction of relevant data elements
5. Appraise – critical appraisal of full-text articles, including bias risk assessment

Table 1. AI terminology relevant for systematic literature reviews

Term	Definition
Machine learning	An application of AI that enables computer systems to learn and improve from experience (typically from large amounts of training data) without being explicitly programmed
Natural language processing	A component of AI that applies computational techniques to analyse human language as it is spoken or written
Text classification	Automated categorisation of documents into groups of interest
Data extraction	The task of identifying key data elements and information from texts (e.g., study population, outcomes)
Semi-automation	Using machine learning to increase the speed and efficiency of review tasks rather than to execute them autonomously
Human-in-the-loop	Workflows in which humans remain involved and are supported, rather than replaced, by AI (i.e., semi-automation)

Source: Adapted from Marshall and Wallace.[6]

6. Analyse – qualitative and quantitative data analysis, including meta-analysis, where appropriate

The conventional approach to systematic reviews requires highly skilled resources for what are mostly manual and repetitive tasks. Searches need to be set up and run in multiple databases. Screening and appraisal tasks are generally duplicated by two reviewers with disagreements resolved by a third reviewer. And depending on the number of relevant articles, data extraction is time-consuming and requires additional quality checks. While there are commercial packages available to aid a collaborative review process, many reviewers still rely on basic spreadsheets and reference manager software to track and document their reviews. For EU MDR compliant reviews, each review step also needs to be sufficiently documented so that they can be reproduced for future updates. For companies with many medical devices marketed in Europe, maintenance of reviews for each product to meet regulatory requirements becomes quickly untenable.

AI-supported systematic reviews approach

The steps to undertake an AI-assisted systematic are the same as for conventional reviews. The key difference between the two approaches is the extent to which individual steps in the process could be automated, or rather semi-automated (see Table 1) using AI technology. Most of the labour-intensive review tasks – screening, data extraction, and to some extent critical appraisal – could be supported by AI-based tools. Table 2 provides some examples of AI-based tools already available that can be used to support distinct tasks of systematic reviews.

Most of these tools use AI to support just one discrete task in the overall review process, and most employ a "human-in-the-loop" workflow, in that they do not intend to replace human reviewers, but rather to make the reviewer more efficient.

Screening

The screening step of the review process is one of the most time-consuming, with much potential for optimisation through the use of AI. This is also an area where AI research efforts have been concentrated, with some tools mature enough to be implemented in your next systematic reviews. After removing duplicate publications from the search results, screeners may have to read several hundred abstracts and quickly and accurately determine if the abstract meets the inclusion requirements of the review. Machine learning tools used for the screening process are designed to learn from the decision of the human reviewer whether to include or exclude each reference reviewed. As the system learns, references are continuously prioritised and sorted by their likelihood for inclusion. This can focus the screener on the records most likely to meet the inclusion criteria and potentially speed up the entire review process. The potential time savings that could be gained by priority ranking references using machine learning have been demonstrated in a user study of the screening tool RobotAnalyst.[7]

> The conventional approach to systematic reviews requires highly skilled resources for what are mostly manual and repetitive tasks.

AI-based screening tools can also serve as a second screener. Another systematic review package familiar to some medical device writers is DistillerSR, a web-based reference screening, data extraction and reporting solution for systematic reviews. Since 2018, DistillerSR has used AI-supported reference screening that uses machine learning and natural language processing. There are several other examples of AI tools that support the screening process using machine learning, including Rayyan and SWIFT-review (Table 2).

Data extraction

Data extraction is another systematic review task where AI applications are showing promise. The level of extraction provided by each tool can vary from identifying and highlighting sentences that are deemed most likely to contain relevant information to extraction of a specific data element. RobotReviewer is one example of such

Table 2. Examples of AI tools intended to support systematic review tasks

Systematic review task	AI-based functionality
Screening	
DistillerSR (https://www.evidencepartners.com/) Rayyan[8] (https://rayyan.qcri.org/) RobotAnalyst[6] (http://www.nactem.ac.uk/robotanalyst/) SWIFT-Review[9] (https://www.sciome.com/swift-review/)	Screening systems automatically sort a search retrieval by relevance (probability-based) determined using machine learning and text mining functionality. The relevancy predictions are continuously updated based on the reviewer's prior selections. Some tools also utilise topic modelling where related abstracts are automatically grouped.
Date extraction	
ExaCT[10] (http://exactdemo.iit.nrc.ca/) RobotReviewer (https://www.robotreviewer.net)	These prototype systems automatically extract key data elements, such as PICO information and sample size, from unstructured written text. Automatically extracted content is presented through a web-based interface that assists the human reviewer in verifying and changing the extracted information for each data element.
Appraisal (risk of bias assessment)	
RobotReviewer[11] (https://www.robotreviewer.net)	This tool attempts to detect risk of bias in randomised controlled trials using a machine learning algorithm; the human reviewer confirming the initial assessment of the tool (semi-automation). A free demo tool is available at https://robotreviewer.vortext.systems/.

Abbreviations: PICO, population, intervention, comparator, outcome.

This list is not intended to be exhaustive; see SR Toolbox[12] for more complete and up-to-date lists.

a tool that can extract text describing population, intervention, control, and outcomes – the so-called PICO elements. For a fun demonstration, take a PDF of your favourite randomised controlled trial publication and drag and drop the file into their demo tool available at https://robotreviewer.vortext.systems/. It will automatically highlight text throughout the document that describes each PICO element (Figure 1) and generates a report that summarises the study characteristics and main findings. One limitation of this tool is that it captures both intervention and control together under the single "intervention" label. A human reviewer is still needed to interpret the intervention under study and the control treatment.

In a systematic review of methods used for data extraction for systematic review, the authors found that many methods aimed to extract relatively straightforward data elements; the most frequently studied data elements were participant characteristics, interventions, and outcomes (as seen with the RobotReviewer example).[13] Many other important data elements, such as duration of follow-up or incidence of adverse events for each participant group, have been studied to a lesser extent or not at all. Another limitation of the current AI tools is that most are limited to evaluation of randomised controlled trials. This is a barrier for adoption by the medical device writer as a substantial amount of the medical device data used in clinical evaluations comes from observational studies. Some research groups have acknowledged this gap and are working to expand the body of AI research in this area.[14]

Critical appraisal

AI is also being used to appraise publications selected for full-text review, for example by assessing the risk of bias. One example system attempting this task is again RobotReviewer.[14,15] The tool analyses the full text of a publication and identifies information about randomisation, allocation concealment, and blinding of participants and outcome assessment to generate a bias assessment report using the Cochrane risk of bias tool (Figure 2). In addition to the bias assessment report, the tool also automatically highlights text relevant to each potential source of bias (analogous to the highlighted PICO elements shown in Figure 1).

Adoption of AI tools by researchers and industry

AI tools for systematic review have not yet matured sufficiently to see widespread adoption by researchers or industry users. There are many systematic review tools utilising machine learning as an underlying approach, in prototype or early development stage; 17 such tools in the field of healthcare were identified using the SR toolbox search tool.[12] Clear limitations to these systems at this stage include the limited types of data elements that can be extracted and the paucity of research into machine learning applied to observational studies or study designs other than RCTs.[14] Some additional barriers to early adoption are scepticism about the reliability of AI-assisted reviews, a general distrust of handing over an assessment to a computer, or just the logistics of trying to integrate a new tool into an established process.[16] While most of the example tools in Table 2 are either completely free or offer free versions, the cost of some commercial tools may be prohibitively expensive, especially for the freelance medical device writer.

A quick informal survey of members of EMWA's medical device special interest group revealed that few had experience with AI-based tools. DistillerSR and Rayyan were mentioned, but not used regularly, and most writers queried relied on spreadsheets and their preferred reference manager software to carry out reviews. Early adopters of AI-based review tools could

> The full potential of AI-based tools to optimise systematic reviews has not yet been realised, but the field is developing rapidly.

Figure 1. Web-interface of the RobotReviewer demo system showing automatically extracted information on study population, intervention, and outcomes from a PDF publication of a randomised controlled trial

potentially also contribute to their further development by providing data sets of screened and appraised literature that can be used to further train and refine these systems. But it remains to be seen if medical device writers will adopt these tools and be able to successfully integrate them into their review process.

What lies ahead

It seems clear that AI is not about to replace the human systematic reviewer. The full potential of AI-based tools to optimise systematic reviews has not yet been realised, but the field is developing rapidly. If developers can address the limitations of the current tools, such as enabling screening of study designs other than RCTs and expanding the possibilities for data extraction, their appeal to the medical writer will grow. Companies and writers that are tasked with the creation and maintenance of clinical evaluation reports should evaluate the potential advantages of adopting some of these emerging tools into their clinical evaluation processes. A validated, reliable, and easy-to-use tool that incorporates AI technology to support multiple steps of the systematic review process for multiple study designs is hopefully not too far away. Medical device writers should be on the lookout for such tools that could optimise the systematic review process as this exciting field continues to develop.

Trial	Design				
Dishman RK, 2010	RCT	?	?	?	?
Fjeldsoe BS, 2010	RCT	+	?	?	?
Online R, 2008	RCT	?	?	?	?
Furber S, 2010	RCT	+	+	?	?
		Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of output assessment

Figure 2. Example bias assessment table generated by RobotReviewer from full-text analysis of four RCTs

Acknowledgements

The author would like to thank Dr Byron Wallace and the developers of RobotReviewer for providing screenshots from their demo system. The author also thanks Beatrix Doerr for her review and input to the draft manuscript.

Disclaimers

The opinions expressed in this article are the author's own and are not necessarily shared by her employer or EMWA.

Conflicts of interest

The author declares no conflicts of interest.

References

1. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. [cited 9 Sep 2019]. Available from: https://eur-lex.europa.eu/eli/reg/2017/745/2017-05-05.
2. Garner P, Hopewell S, Chandler J, MacLehose H, Schunemann HJ, Akl EA, et al. When and how to update systematic reviews: consensus and checklist. BMJ (Clinical research ed). 2016;354:i3507.
3. Policy: Cochrane Review updates 2019 [cited 7 Sep 2019]. Available from: https://documentation.cochrane.org/display/EPPR/Policy%3A+Cochrane+Review+updates.
4. Extance A. How AI technology can tame the scientific literature. Nature. 2018;561(7722):273–4.
5. Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]. The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org.
6. Marshall IJ, Wallace BC. Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. Syst Rev. 2019;8(1):163.
7. Przybyla P, Brockmeier AJ, Kontonatsios G, Le Pogam MA, McNaught J, von Elm E, et al. Prioritising references for systematic reviews with RobotAnalyst: A user study. Res Synth Methods. 2018;9(3):470–88.
8. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev. 2016;5(1):210.
9. Howard BE, Phillips J, Miller K, Tandon A, Mav D, Shah MR, et al. SWIFT-Review: a text-mining workbench for systematic review. Syst Rev. 2016;5(1):87.
10. Kiritchenko S, de Bruijn B, Carini S, Martin J, Sim I. ExaCT: automatic extraction of clinical trial characteristics from journal publications. BMC Med Inform Decis Mak. 2010;10:56.
11. 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.
12. SR Tool Box. 2019 [cited 9 Sep 2019]. Available from: http://systematicreviewtools.com/index. php.
13. Jonnalagadda SR, Goyal P, Huffman MD. Automating data extraction in systematic reviews: a systematic review. Syst Rev. 2015;4:78.
14. Karystianis G, Thayer K, Wolfe M, Tsafnat G. Evaluation of a rule-based method for epidemiological document classification towards the automation of systematic reviews. J Biomed Inform. 2017;70:27–34.
15. Soboczenski F, Trikalinos TA, Kuiper J, Bias RG, Wallace BC, Marshall IJ. Machine learning to help researchers evaluate biases in clinical trials: a prospective, randomized user study. BMC Med Inform Decis Mak. 2019;19(1):96.
16. O'Connor AM, Tsafnat G, Thomas J, Glasziou P, Gilbert SB, Hutton B. A question of trust: can we build an evidence base to gain trust in systematic review automation technologies? Syst Rev. 2019;8(1):143.

Author information

Kelly Goodwin Burri has 15 years of experience in bioengineering research, clinical development, and epidemiology. She currently works as a medical writer for Stryker GmbH preparing clinical evaluation reports and related documentation for medical devices in the fields of orthopaedics and traumatology.

EMWA NEEDS YOU

EMWA is a member-run organisation

When you volunteer to assist EMWA in any capacity, you are furthering the development of our association.

You can choose how you want to get involved: in a very limited way or as part of a larger project. The choice is yours, and everyone shares the benefits.

EMWA members can volunteer in the following areas:

Conference
- Planning committee
- Advertising

Finance

Journal
- Contributions
- Editorial Board

Website
- Contributions
- Web Team

Freelance Business Group

Social Media Team

Training
- Leading workshops
- Professional Development Programme Committee
- Webinar contributions
- Webinar Team

Internship project

Special Interest Groups
- Pharmacovigilance
- Regulatory Disclosure
- Medical Devices
- Ambassador programme

Executive Committee
- President
- Vice President
- Journal Editor
- Public Relations Chair
- Conference Chair
- Honorary Secretary
- Professional Development Programme Committee Chair
- Treasurer

TO FIND OUT MORE

If you are a member of EMWA and eager to support ongoing initiatives, please contact info@emwa.org.

WHY VOLUNTEER?
- Help promote the role of medical writers and strengthen our association
- Help to raise standards in our field
- Increase your visibility and communication opportunities within the medical writing community
- Add some prestige to your CV
- Improve your knowledge of medical writing and related topics