



Automation in Healthcare Systematic Review

Raquel L. Ruiz^(✉) and Vincent G. Duffy^(✉)

Purdue University, West Lafayette, IN 47906, USA
{ruiz142, duffy}@purdue.edu

Abstract. Systematic reviews are an essential tool to help healthcare providers and medical practitioners stay up to date on the latest evidence and practices within their field. In recent years, developments have arisen in automating these systematic reviews. While automated systematic reviews are not currently in widespread use in the medical field, they present a way to increase the output of systematic reviews. This study is a literature review of how the automating of systematic reviews is currently being integrated into the healthcare system. This literature review will be relying on tools such as Web of Science, Harzing’s Publish or Perish, VOSviewer, CiteSpace, and MAXQDA to collect and analyze article data sets based on keywords “automating systematic reviews” and “healthcare systematic reviews”. Co-citation and content analyses were performed on data sets to determine which articles were most relevant. These analyses showed that automating systematic reviews is developing more within fields outside of the medical field. However, researchers are beginning to take an interest in the applications of an automated systematic review to use in the medical field. The healthcare sector has been more hesitant to adopt automation and often continues to rely on traditional methods. Automation technology is quickly advancing but currently lacks the ability to apply critical thinking of how literature is relevant to current practice, more easily understood by clinician experience, attitudes, and values.

Keywords: Systematic review · Healthcare · Automation · Bibliometric analysis · Harzing · VOSviewer · CiteSpace · MAXQDA

1 Introduction and Background

1.1 Systematic Reviews in Healthcare

A systematic review is a summary of primary research on a specific topic to collect information and reduce bias. It “typically involve[s] a detailed and comprehensive plan and search strategy”, “reducing bias by identifying, appraising, and synthesizing all relevant studies” [1]. With ever-increasing advancements in medical technology, practices, and knowledge, medical practitioners need to be well informed of the most recent evidence. Systematic reviews offer medical practitioners a collection of primary studies to assist them with decision-making. Not only are systematic reviews useful for obtaining up-to-date information, but they are also used when developing new clinical practice guidelines [2]. Systematic reviews are relied upon for the prevention of errors, patient satisfaction and safety, improving the quality of care, and saving lives. Deaths from

healthcare errors are estimated to be over 250,000 a year, with many likely unreported. Based on these numbers, healthcare errors are the 3rd leading cause of death in the United States [3]. Challenges arise for busy healthcare clinicians to fully access information of similar and relatable trials which may be applicable to quality and safety improvements specific to their areas of practice. Reviews are rapidly added, making it difficult to keep up to date [4]. Many of these reviews may not be fully updated or amended when new data arises. As healthcare advancements rapidly change, healthcare clinicians have less time to research, evaluate, educate, and adopt new evidence-based practices. Inadequate reporting can create poor interpretation and critical analysis of systematic reviews. While they have many drawbacks, they are still more advanced in comparison to traditional “opinion reviews that cite evidence selectively to support a viewpoint” [5]. Finally, concerns with validity exist within these trials as limitations and bias, whether transparent or unrecognized, along with sample populations and targeted responses to questions may lack relevance to current clinical practice. Systematic reviews require extensive evaluation and time to increase the likelihood of being applicable to current practice [5].

1.2 Systematic Review Automation

With evidence-based publications arising at rapid rates, healthcare guidelines are continuously requiring revisions to be up to date. Significant delays in literature review for application in clinical practice make achieving improved patient outcomes difficult to achieve. A proposed solution is developing and adopting automation tools to search, gather, and more effectively and efficiently group and analyze publications [4]. Automation can serve by increasing the speed of data acquisition and generate valid, up-to-date recommendations for best practices. Using artificial intelligence within automation can serve to solve complex problems and accelerate advancements in health-related practice and treatments [6]. In addition, automation may reduce the risk of bias in the grouping of publications or omission of other relevant publications that do not serve the interests of the reviewer. While automation may lead to greater efficiency, concerns arise with “the extent to which an innovation is in line with current values and practices” [4]. Other challenges with automation include the need for high relatability and transparency which requires critical thinking and experience [7].

2 Purpose of Study

The purpose of this study is to conduct a literature review on how automation is being integrated into the healthcare systematic review process. Automated data collection has been gaining attention from both large and small businesses in the past decade. While it has made progress in the healthcare sector, there is still hesitation to fully implement an automated systematic review for medical practices. This study will look at how automating healthcare systematic reviews are perceived by the medical community and understand why there is hesitation in fully automating this process within the medical field. This study uses programs such as Web of Science, Harzing’s Publish or Perish, VOSviewer, CiteSpace, and MAXQDA. The program Zotero was used to create the citations and reference list [8].

3 Methodology

3.1 Data Collection

Data was initially collected from doing a keyword search within the Web of Science database. This search was later refined with a larger data pool using the program Harzing’s Publish or Perish to conduct a keyword search in both the Web of Science and Google Scholar databases. Both databases were used since each has its advantages and disadvantages. Web of Science offers data on the title, authors, source, abstract, and cited references, but only allows 500 articles per data set. Google Scholar does not include data on cited references but offers up to 1000 articles in a data set from a larger amount of articles. Since cited references are required for co-citation analysis, the data from Google Scholar cannot be used. However, the larger data set from Google Scholar is preferable for cluster analysis. The keywords used in the data collection were “automating systematic reviews” and “healthcare systematic reviews”.

3.2 Trend Analysis

The trend analysis was conducted using data from the Web of Science along with analysis tools provided within the program. All data was collected between the years 1980 and 2021.

Figure 1 shows the number of articles published per year that were found using a keyword search of “automating systematic reviews”. This graph shows that interest in this topic first began in 1993, but gained interest in the 2010s. The year 2021 shows a significant decrease in article publication due to the year not being completed. Based on the increasing trend of prior years, it can be predicted that the year 2021 will have an even greater amount of published articles than the years prior.

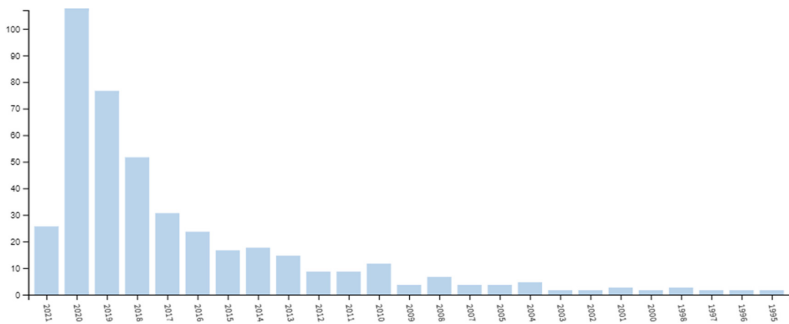


Fig. 1. Trend analysis of data from the articles on automating systematic reviews [9]

Figure 2 is created with the same data as Fig. 1, but it shows how many articles are written in each field. This graph shows that most articles that have the keyword “automating systematic review” are from the computer science information systems field. There is a significant decrease in articles of over 20 articles between the computer

science information systems field and the remaining fields. This graph also shows the field of health care sciences services is ranked 7th for this keyword search with 24 articles. The information from Fig. 2 can suggest that most of the development of automating systematic reviews are occurring within the computer science information systems field and from there, being applied to other fields.

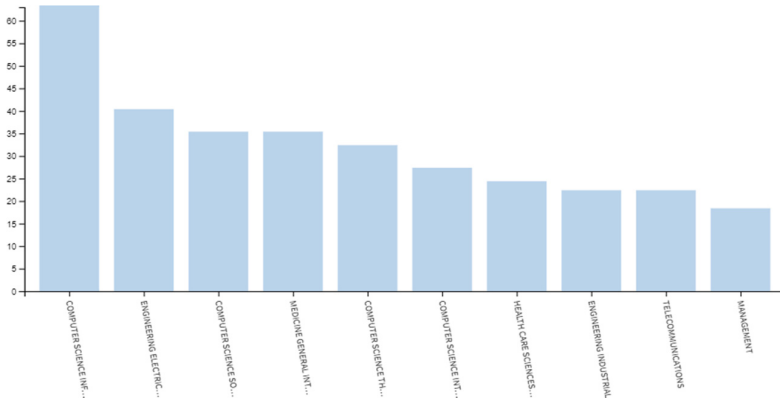


Fig. 2. Trend analysis of data from the articles on automating systematic reviews organized by field [9]

Figure 3 is created with the same data as Fig. 1 and Fig. 2 but has refined the keyword search to only include articles within the medical field. This decreased the articles in the data set from 416 articles to 35 articles. This graph shows the year 2019 was the peak of when articles with the keyword “automating systematic review” were published. This contradicts the trend of increasing published articles in Fig. 1. The lower amount of published articles in 2020 could have been due to resources being put towards the COVID-19 pandemic.

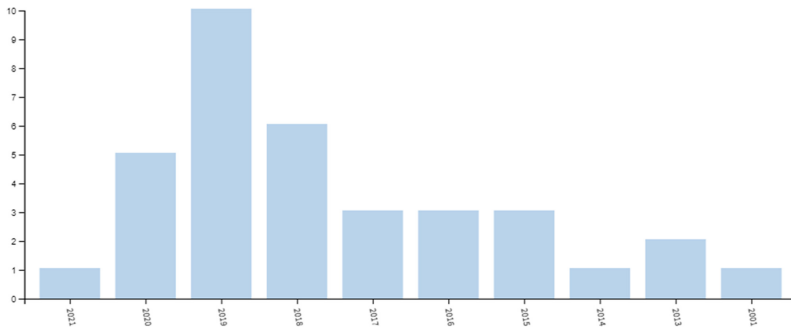


Fig. 3. Trend analysis of data from the articles on automation of systematic reviews in the medical field [9]

Figure 4 is a graph of the number of articles published per year that were found using a keyword search of “healthcare systematic reviews”. This graph shows a trend of an increase in published articles. The year 2020 had the largest amount of published articles. This lowers the likeliness of the explanation for the decrease in articles in the year 2020 for Fig. 3 to be true.

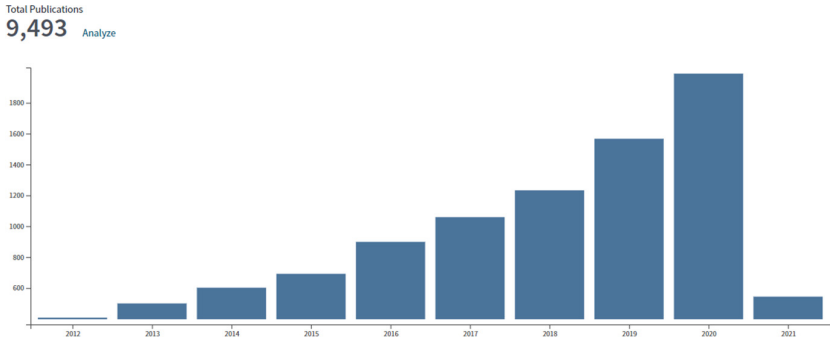


Fig. 4. Trend analysis of data from the articles on healthcare systematic reviews [9]

Figure 5 is a pivot table showing the results of an analysis of the sum of articles per author. Figure 5 was created from data collected using a keyword search of “automating systematic reviews in healthcare” conducted with Harzing’s Publish or Perish from the Google Scholar database. A pivot table is a tool that can convert 65,000 rows of data into a summarized table [10]. The author with the largest number of articles is James Thomas, with nine articles. Most of the articles he has published are on the topic of automation and computer assistance. While not directly in the medical field, his articles show interesting perspectives on automation. One such article is “On the use of computer-assistance to facilitate systematic mapping” [11]. The author with the second largest amount of published articles is Paul Glasziou who began working on the topic of automating systematic reviews when the topic gained popularity. One of his earlier publications is “The automation of systematic reviews” [12].

4 Results

4.1 Co-citation Analysis

The co-citation analysis was conducted with the program VOSviewer with data extracted from the article search in the Web of Science. Specifically, the data used to make Fig. 3 was used to create Fig. 5. The exported data included the title, authors, source, abstract, and cited references. Co-citation analyses are used as a way to adequately identify the relationships and structure between research areas [15]. The basis of the co-citation analysis is the higher amount of co-citations between authors there is, the stronger the relationship of the authors is [16]. The VOSviewer creates a cluster map with the imported data that shows the articles that we cited the most. The links between the authors show

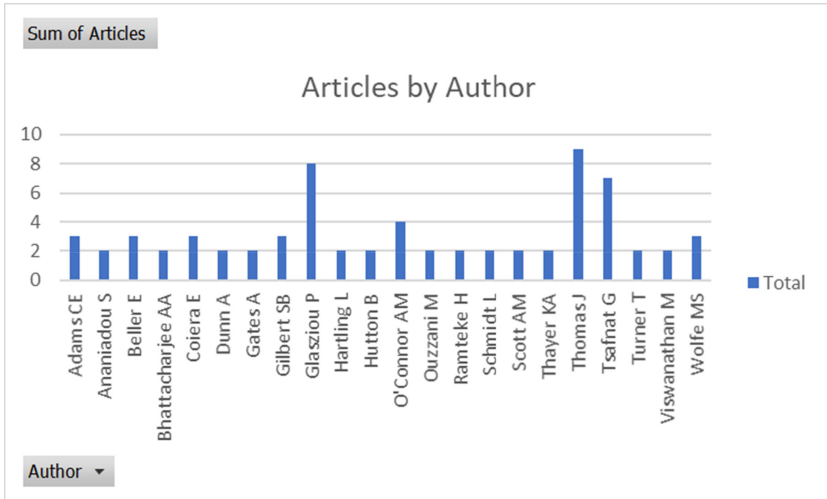


Fig. 5. Pivot table in Excel of the number of articles per author [13, 14]

which articles have been cited together. Limits can be set for the minimum number of citations for each article. Another program that can perform a co-citation analysis is CiteSpace. This program uses the same data as the VOSviewer, but a unique output it has is a citation burst. A citation burst is a way that makes “it possible to find the articles that receive particular attention from the related scientific communities in a certain period of time” [17]. Not only is it able to perform co-citation analyses, but it is also able to assist in “identifying betweenness centrality between pivotal points in scientific articles, which indicates the significance of the nodes in a network” [17]. CiteSpace is able to identify co-cited articles by choosing different nodes [17].

VOSviewer Analysis. Figure 6 is a co-citation cluster map of the data from the keyword search for “automating systematic reviews” within the medical field. The minimum number of co-citations was set to 10. Figure 6 contains 2 clusters, 44 links, and has a total link strength of 244. The article with the largest citations was “Using text mining for study identification in systematic reviews: a systematic review of current approaches” [18] with 22 citations. The article with the second largest citations was “Systematic review automation technologies” [19] with 21 citations. The earliest article within this cluster is from 2009, with most of the other articles from the mid to late 2010’s. This appears to follow the article publication trend from Fig. 4.

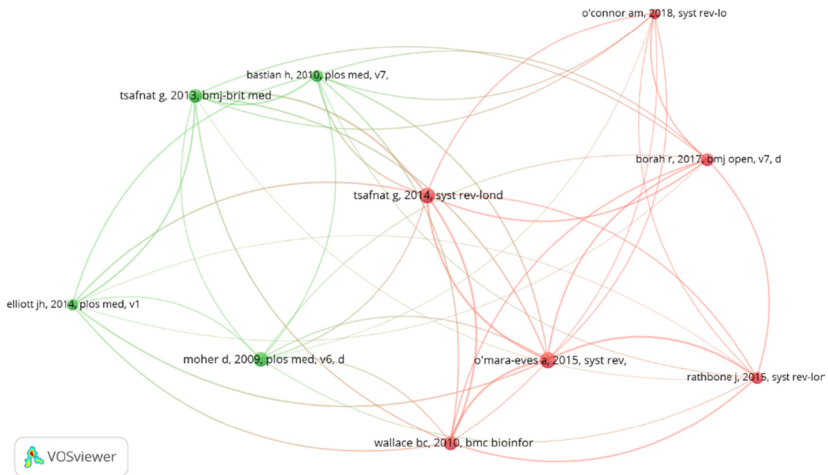


Fig. 6. Co-citation analysis for automating systematic reviews in the medical field [20]

CiteSpace Analysis. Figure 7 is a cluster burst created with CiteSpace using the data from the keyword search “healthcare systematic reviews”. There are 5 bursts within Fig. 7 that show the main keyword per cluster.

Figure 8 is a list of references from the data in Fig. 7 with the strongest citation burst. The top reference is “Grading Quality of Evidence and Strength of Recommendations: A Perspective” [21].

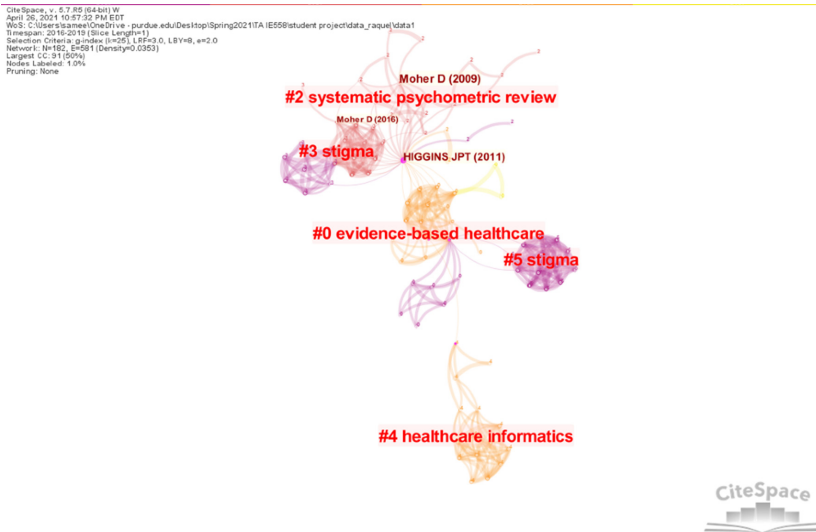


Fig. 7. Citation burst from CiteSpace on data for healthcare systematic reviews [22]

Top 1 References with the Strongest Citation Bursts


References	Year	Strength	Begin	End	2016 - 2019
Moher D, 2009, PLOS MED, V6, P0	2009	4.74	2016	2017	

Fig. 8. References from Fig. 8 with the strongest citation bursts [22]

Figure 9 is a cluster burst created with CiteSpace using the data from the keyword search for “automating systematic reviews” within the medical field.

Figure 10 is a list of references from the data in Fig. 9 with the strongest citation burst. The top reference happens to be the same as Fig. 8, “Grading Quality of Evidence and Strength of Recommendations: A Perspective” [21]. The second top reference is “Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up?” [23]. Figure 8 and Fig. 10 show that while there has been a significant increase in articles written for automating systematic reviews for healthcare, the most referenced article is from 2009. Technology has increased significantly since this paper was published; however, Fig. 8 and Fig. 10 suggest that there is still a lack of research and integration on this topic within the medical field.



Fig. 9. Citation burst from CiteSpace of data for automating systematic reviews in the medical field [22]

Top 2 References with the Strongest Citation Bursts



References	Year	Strength	Begin	End	2016 - 2019
Moher D, 2009, PLOS MED, V6, P0	2009	1.87	2016	2017	
Bastian H, 2010, PLOS MED, V7, P0, DOI 10.1371/journal.pmed.1000326, DOI	2010	1.8	2017	2019	

Fig. 10. References from Fig. 8 with the strongest citation bursts [22]

4.2 Co-authorship Analysis

A co-authorship analysis is used as “a proxy of research collaboration” “that links different sets of talent to produce a research output” [24]. It is a useful analysis to understand where articles are coming from and what social networks have an interest in specific research topics. Co-authorship also allows the “relations between these groups and changes in time” between these groups to be analyzed [25].

Figure 11 is a cluster map of a co-authorship analysis for the data from the keyword search for “automating systematic reviews” within the medical field. Figure 11 contains 4 clusters, 41 links, and has a total link strength of 84. The authors with the most connections are James Thomas, Paul Glasziou, and Guy Tsafnat. These results are confirmed with the trend analysis in Fig. 5 since these authors have the most published articles so they will most likely have the largest co-authorship within this keyword search. Together Fig. 4, Fig. 5, Fig. 11 suggest that more authors have taken interest and begun to publish papers.

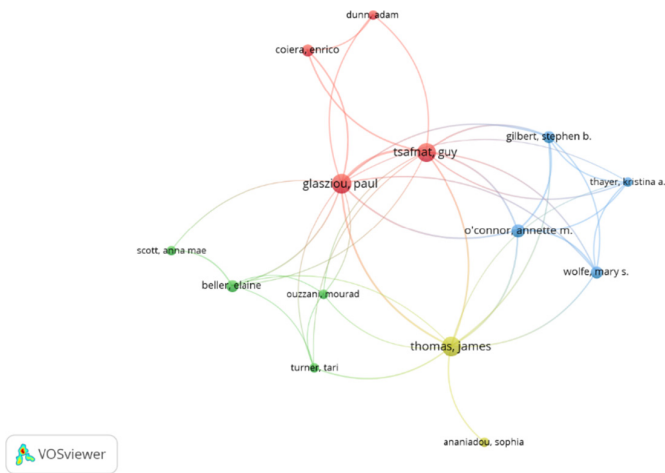


Fig. 11. Co-authorship analysis on data for automating systematic reviews in the medical field [20]

4.3 Content Analysis

Comparable with co-citation analysis, content analysis is conducted with the program VOSviewer. However, it uses data collected from Google Scholar using the program Harzing’s Public or Perish [13]. Since content analysis does not require cited references in the data, data from Google Scholar can be used. The advantage to using data from Google Scholar is the ability to export a larger data pool compared to Web of Science. Content analyses are used to organize and interpret the data from imported articles to gain reasonable results and understanding from the set of articles [26].

Figure 12 is a content analysis cluster map produced from the data collected from the keyword search for “systematic review in healthcare”. The data set was collected

from Google Scholar and contains 980 articles. The minimum occurrence was set to 20 which resulted in 51 keywords but only the 60% most relevant keywords which is 31. Figure 12 contains 5 clusters, 287 links, and has a link strength of 1918. The keywords with the largest connection are “systematic review” with 692 occurrences, ‘healthcare’ with 177 occurrences, and “systematic literature review” with 121 occurrences.

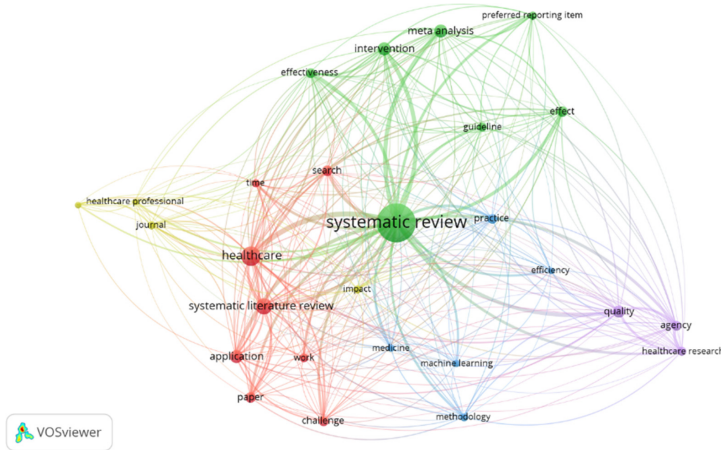


Fig. 12. Cluster analysis on data for automating systematic review in healthcare [20]

Figure 13 is a content analysis cluster map produced from the data collected from the keyword search for “automating systematic review” in the medical field. This data set was exported from the Web of Science and contains 35 articles. Figure 9 has 3 clusters, 78 links, and a total link strength of 1112. The minimum number of articles was set to 20 occurrences and 16 keywords met this condition. The top keywords from this content analysis cluster are systematic review, automation, and study. These keywords had occurrences of 84, 83, and 72, respectively. Both Fig. 12 and Fig. 13 show that the articles within the data sets used focus on automation and systematic reviews.

4.4 Content Analysis from MAXQDA

MAXQDA performs a similar content analysis as VOSviewer, however, MAXQDA uses a smaller set of data and is able to examine this data in greater detail. MAXQDA also allows the user to manually remove stop words from being analyzed within the data. MAXQDA is also able to visualize results into a word cloud, which can be an effective way to communicate the most common words within articles [27].

Figure 14 is a word cloud created from a data set collected from Web of Science using a keyword search of “healthcare automation”. This word cloud shows that the top three keywords found in the articles in the data set are “systematic”, “automation”, and “review”. Figure 14 shows that recent articles having to do with “healthcare automation” are beginning to focus on systematic reviews.

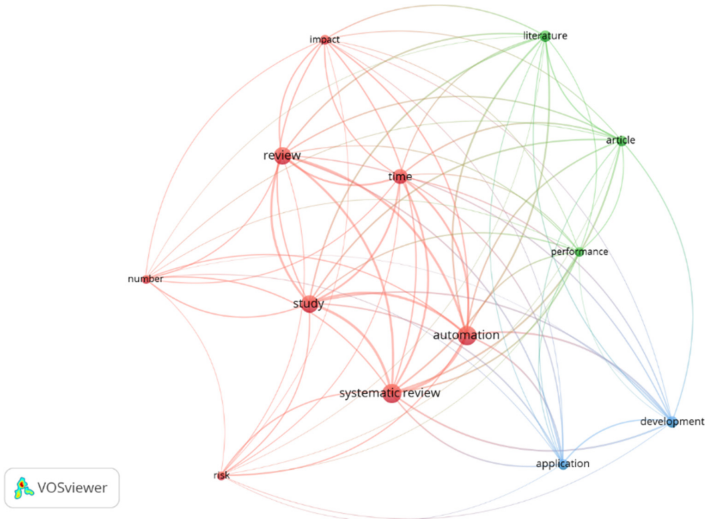


Fig. 13. Cluster analysis on data for automating systematic reviews in the medical field [20]



Fig. 14. Word Cloud from MAXQDA [28]

5 Conclusion

A systematic review of literature is critical for the advancement of healthcare practice, safety, and improved patient outcomes. The healthcare sector has been more hesitant to adopt automation and often continues to rely on traditional methods. Challenges exist in both traditional and advanced technology methods. Shortcomings with traditional methods include significant delays in the collection of relevant and reliable data, a potential bias that may result in omitting or overuse of selected articles and keeping up to date with the rapid additions of research literature. Automation technology is quickly

advancing but currently lacks the ability to apply critical thinking of how literature is relevant to current practice, more easily understood by clinician experience, attitudes, and values. Developing and utilizing further advancements in technology, including artificial intelligence may lead to maximizing the full potential of systematic literature reviews. Understanding and addressing concerns of healthcare clinicians such as ensuring technology is well-matched with current practice and values, is vital for acceptance and trust of automation.

6 Future Work

Advancements in artificial intelligence, natural language processing, and machine learning have created more advanced algorithms that can enhance the efficiency and quality of automated systematic reviews. This complex ecosystem can provide interconnections of numerous techniques in artificial intelligence to solve complex problems within the automation process. This may lead to faster activity, more accurate, and less expensive results. Machine learning is a promising type of artificial intelligence that may accelerate quality outcomes [29]. Funding challenges have led to delays and fragmented development. Collaboration of automated tools are critical towards overcoming this barrier. In 2015, the International Collaboration for the Automation of Systematic Reviews was founded to exchange data and develop more efficient toolkits. Members include a wide variety of interdisciplinary professions such as software engineers, librarians, statisticians, experts in artificial intelligence, linguists, and researchers [29]. The NSF is also currently funding the University of Arkansas System with an award for collaborative research on automated knowledge discovery in reliability and healthcare from complex data with covariates [30]. This award has already helped produce publications such as “Flexible methods for reliability estimation using aggregate failure-time data” [31] and “Comparison study on general methods for modeling lifetime data with covariates” [32].

References

1. Uman, L.S.: Systematic reviews and meta-analyses. *J. Can. Acad. Child Adolesc. Psychiatry* **20**(1), 57–59 (2011)
2. Gopalakrishnan, S., Ganeshkumar, P.: Systematic reviews and meta-analysis: understanding the best evidence in primary healthcare. *J. Fam. Med. Prim. Care* **2**(1), 9–14 (2013). <https://doi.org/10.4103/2249-4863.109934>
3. Anderson, J.G., Abrahamson, K.: Your health care may kill you: medical errors. *Stud. Health Technol. Inform.* **234**, 13–17 (2017)
4. Arno, A., Elliott, J., Wallace, B., Turner, T., Thomas, J.: The views of health guideline developers on the use of automation in health evidence synthesis. *Syst. Rev.* **10**(1), 16 (2021). <https://doi.org/10.1186/rs.3.rs-23742/v2>
5. Maggio, L.A., Sewell, J.L., Artino, A.R., Jr.: The literature review: a foundation for high-quality medical education research. *J. Grad. Med. Educ.* **8**(3), 297–303 (2016). <https://doi.org/10.4300/JGME-D-16-00175.1>
6. O’Blenis, P.: Past, Present, and Future: Automation in Systematic Review Software. <https://blog.evidencepartners.com/past-present-and-future-automation-in-systematic-review-software>. Accessed 25 Apr 2021

7. Marshall, I.J., Wallace, B.C.: Toward systematic review automation: a practical guide to using machine learning tools in research synthesis. *Syst. Rev.* **8**(1), 163 (2019). <https://doi.org/10.1186/s13643-019-1074-9>
8. Zotero. Corporation for Digital Scholarship. <https://www.zotero.org/>
9. Web of Science. https://apps-whoofknowledgecom.ezproxy.lib.pur-due.edu/WOS_GeneralSearch_input.do?product=WOS&search_mode=General-Search&SID=7EKUw7yEVcCGVrUwOlu&preferencesSaved=
10. Alexander, M., Jelen, B.: *Pivot Table Data Crunching*. Pearson Education (2001)
11. Haddaway, N.R., et al.: On the use of computer-assistance to facilitate systematic mapping. *Campbell Syst. Rev.* **16**(4), e1129 (2020). <https://doi.org/10.1002/cl2.1129>
12. Tsafnat, G., Dunn, A., Glasziou, P., Coiera, E.: The automation of systematic reviews. *BMJ* **346**, f139 (2013). <https://doi.org/10.1136/bmj.f139>
13. Anne-Wil, H.: Harzing's Publish or Perish. <https://harzing.com/resources/publish-or-perish>
14. Persson, O.: BibExcel. <https://sites.google.com/site/bibexcel2015/>
15. Jeong, Y.K., Song, M., Ding, Y.: Content-based author co-citation analysis. *J. Informetr.* **8**(1), 197–211 (2014). <https://doi.org/10.1016/j.joi.2013.12.001>
16. Andrews, J.E.: An author co-citation analysis of medical informatics. *J. Med. Libr. Assoc.* **91**(1), 47–56 (2003)
17. Zhou, W., Chen, J., Huang, Y.: Co-citation analysis and burst detection on financial bubbles with scientometrics approach. *Econ. Res.-Ekon. Istraživanja* **32**, 2310–2328 (2019). <https://doi.org/10.1080/1331677X.2019.1645716>
18. O'Mara-Eves, A., Thomas, J., McNaught, J., Miwa, M., Ananiadou, S.: Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Syst. Rev.* **4**(1), 5 (2015). <https://doi.org/10.1186/2046-4053-4-5>
19. Jan van Eck, N., Waltman, L.: VOSviewer. Leiden University's Centre for Science and Technology Studies. <https://www.vosviewer.com/>
20. Ansari, M.T., Tsertsvadze, A., Moher, D.: Grading quality of evidence and strength of recommendations: a perspective. *PLOS Med.* **6**(9), e1000151 (2009). <https://doi.org/10.1371/journal.pmed.1000151>
21. Chen, C.: CiteSpace. Drexel University. <http://cluster.cis.drexel.edu/~cchen/citespace/>
22. Bastian, H., Glasziou, P., Chalmers, I.: Seventy-five trials and eleven systematic reviews a day: how will we ever keep up? *PLOS Med.* **7**(9), e1000326 (2010). <https://doi.org/10.1371/journal.pmed.1000326>
23. Kumar, S.: Co-authorship networks: A review of the literature. *Aslib J. Inf. Manag.* **67**, 55–73 (2015). <https://doi.org/10.1108/AJIM-09-2014-0116>
24. Peters, H.P.F., Van Raan, A.F.J.: Structuring scientific activities by co-author analysis. *Scientometrics* **20**(1), 235–255 (1991). <https://doi.org/10.1007/BF02018157>
25. Bengtsson, M.: How to plan and perform a qualitative study using content analysis. *NursingPlus Open* **2**, 8–14 (2016). <https://doi.org/10.1016/j.npls.2016.01.001>
26. Lohmann, S., Heimerl, F., Bopp, F., Burch, M., Ertl, T.: Concentri cloud: word cloud visualization for multiple text documents. In: 2015 19th International Conference on Information Visualisation, pp. 114–120 (2015). <https://doi.org/10.1109/iV.2015.30>
27. Kuckartz, U.: MAXQDA. VERBI Software. <https://www.maxqda.com/>
28. Beller, E., et al.: Making progress with the automation of systematic reviews: principles of the International Collaboration for the Automation of Systematic Reviews (ICASR). *Syst. Rev.* **7**(1), 77 (2018). <https://doi.org/10.1186/s13643-018-0740-7>
29. Award Abstract #1635379 Collaborative Research: Automated Knowledge Discovery in Reliability and Healthcare from Complex Data with Covariates. NSF. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1635379&HistoricalAwards=false

30. Karimi, S., Liao, H., Fan, N.: Flexible methods for reliability estimation using aggregate failure-time data. *IISE Trans.* **53**(1), 101–115 (2021). <https://doi.org/10.1080/24725854.2020.1746869>
31. Liao, H., Karimi, S.: Comparison study on general methods for modeling lifetime data with covariates. In: *2017 Prognostics and System Health Management Conference (PHM-Harbin)*, Harbin, China, July 2017, pp. 1–5 (2017). <https://doi.org/10.1109/PHM.2017.8079122>