Caring is Sharing – Exploiting the Value in Data for Health and Innovation
M. Hägglund et al. (Eds.)

doi:10.3233/SHTI230211

Developing a Comprehensive Search Strategy for the Systematic Review of Clinical Decision Support Systems for Nursing Practice

Cynthia ABI KHALIL[a,b,1], Antoine SAAB[c,b], Jihane RAHME[a] and Brigitte SEROUSSI[b,d]
[a]*Nursing Administration, Lebanese Hospital Geitaoui–UMC, Beirut, Lebanon*
[b]*Sorbonne Universite´,Universite´ Sorbonne Paris Nord, INSERM, Laboratoire d'Informatique Me´dicale et d'Inge´nierie des connaissances en e-Sante´, LIMICS, F-75006 Paris, France*
[c]*Quality and Safety Department, Lebanese Hospital Geitaoui–UMC, Beirut, Lebanon*
[d]*AP-HP, Ho^pital Tenon, Paris, France*
ORCiD ID: Cynthia Abi Khalil https://orcid.org/0000-0001-9112-628X
Antoine Saab https://orcid.org/0000-0003-0673-0875
Brigitte Séroussi https://orcid.org/0000-0002-7676-2832

Abstract. The search strategy of a literature review is of utmost importance as it impacts the validity of its findings. In order to build the best query to guide the literature search on clinical decision support systems applied to nursing clinical practice, we developed an iterative process capitalizing on previous systematic reviews published on similar topics. Three reviews were analyzed relatively to their detection performance. Errors in the choice of keywords and terms used in title and abstract (missing MeSH terms, failure to use common terms), may make relevant articles invisible.

Keywords. Nursing informatics, Clinical decision support systems, Systematic reviews, Precision, Recall

1. Introduction

The literature surrounding the clinical use of clinical decision support systems (CDSSs) in healthcare suggests that these systems are effective to some extent. Published studies usually report on CDSSs targeted on physician performance [1], but the number of studies exploring the success of CDSSs in improving nursing clinical practice has increased in the last years [2], showing there is still room for improvement. In order to develop a CDSS with positive impacts on nursing process, we started with a review of the literature. The search strategy of systematic reviews is of utmost importance, as the quality and validity of the review's findings could be directly affected by the completeness and relevance of identified articles. Errors made in the search process may indeed result in a biased or otherwise incomplete list of articles for the review.

[1] Corresponding Author: Cynthia ABI KHALIL, E-mail: cynthiakjammal@outlook.com

While it is almost impossible, although expected, to build a search strategy that identifies all the relevant scientific articles related to a certain scope, researchers thrive towards having the most comprehensive search results. In practice, this is performed through searching various scientific bibliographic databases (BDBs) such as Embase, Cochrane Library, Web of science, MEDLINE, etc. (the latter being the most popular among researchers in the biomedical sciences field) through a query built using a formal syntax specific to each search engine. While a plethora of scientific and educational resources exist regarding methods to correctly query bibliographic databases, the rate of errors in search strategies of reviews is surprisingly high, reaching 90% with at least half of the errors affecting recall [3]. This also affects Cochrane reviews which is considered as the gold standard [4].

The objective of this work is to analyze the search strategy of published reviews related to CDSSs in nursing practice in order to elaborate an optimized comprehensive search strategy for a systematic review on this topic.

2. Methods

First, a Medline/PubMed query was constructed by two of the authors (CAK and BS), in accordance with the PRESS Guideline Evidence-Based Checklist. The question framework and its delineated concepts were defined, and related terms were harvested so as to build a comprehensive list of conceptual synonyms. The query was built by identifying and extracting the controlled vocabulary terms (i.e., MeSH terms) and in-text synonyms (i.e., “common terms”) that best expressed the research question and the components of the search segments.

As usual, Boolean operator OR was used between terms relative to the same concept in the search segment, so as to increase sensitivity, and Boolean operator AND was used to combine the search segments in order to improve the specificity of retrieved results. Only articles in English, concerning human research, published between 2000 and 2022, with abstract and full-text available were included. The query initially returned 1581 articles among which we found 23 reviews. Reviews with narrow scope (e.g., use of CDSSs in homecare) or specific to one domain or specialty (e.g., use of CDSSs for pressure ulcers) were excluded to only keep eight reviews [1,2,5–10] with a general scope on CDSSs in nursing. These eight reviews all together had selected a total of 138 articles that we used to revisit the query iteratively, adding and removing terms, to gauge the search strategy to include the articles among the 138 articles that were relevant to our search. We finally reached an *optimized* query that returned a total of 3,283 results.

In a second step, the time frames covered by the eight reviews of interest were analyzed (Figure 1). We finally selected three reviews [2,9,10] for evaluation on a common data extraction period (2014-2017), excluding old reviews (more than 10-year-old). The *optimized* query filtered on the same period (2014-2017) yielded 578 articles. Two of the authors (CAK and AS) independently screened all titles, abstracts, and full texts when needed, to classify articles as *relevant*, *non-relevant* or *may-be relevant* (posters, commentaries, short papers were excluded). Consensus between the two researchers was used to resolve any discrepancies. In total, 145 relevant articles published between 2014 and 2017 were finally included to build the “reference list” (RL).

The performance of each of the three reviews to support our search was assessed through the computation of precision defined as the proportion of retrieved articles relevant to the review query, and recall defined as the proportion of relevant articles successfully retrieved by the query.

We adopted a proxy measure based on the comparison of articles retrieved by the reviews and articles included in the RL, assumed as the Gold Standard. We considered that articles retrieved by the reviews were relevant when they were also included in the RL (true positives). Thus, we defined precision as the ratio of the number of articles retrieved by the review *and* included in the RL to the number of articles retrieved by the review, and recall as the ratio of the number of articles retrieved by the review *and* included in the RL to the number of articles included in the RL. To get the number of articles included in the RL but missing in the reviews, authors (CAK, AS, and JR) analyzed the full text version of the 145 RL articles taking into account the scope of each review and the inclusion and exclusion criteria that reviews' authors used.

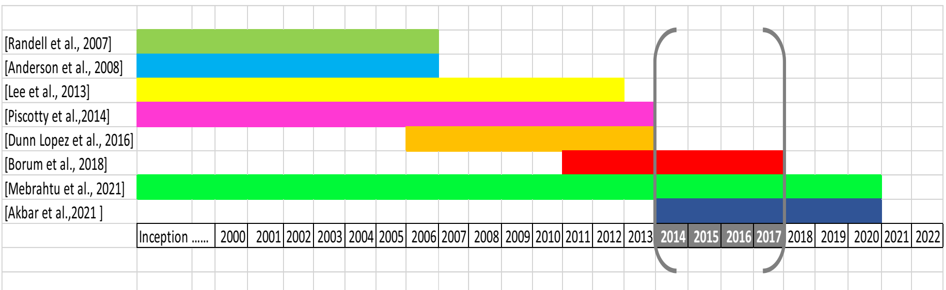

Figure 1. Time frames of reviews to select the largest period that covers the maximum of recent reviews.

3. Results

Akbar et al., [2] retrieved 15 articles in Medline on the [2014;2017] time frame. However, three articles were eliminated because out of the scope of our query (one article had no abstract available, one article was a commentary, and one article was actually a poster) leading to only keep 12 articles for [2]. Thus, the three selected reviews returned 12 [2], 8 [9], and 8 [10] articles in compliance with the inclusion criteria of the optimized query, among which, 11, 5, and 3 articles, respectively, were included in RL.

The optimized query succeeded in retrieving additional 27, 2 and 8 articles published between 2014-2017 and eligible for inclusion for respectively [2], [9] and [10]. Precision on Medline in the selected period is 91.70% (11/12) for [2], 62,50% (5/8) for [9], and 37.50% (3/8) for [10]. In the same way, recall is 28.90% (11/38) for [2], 71.40% (5/7) for [9], and 27.30% (3/11) for [10]. Results are displayed in Table 1.

4. Discussion

Results show that the majority of articles included in reviews are indexed in Medline (15/17 in [2], 8/8 in [9], and 8/9 in [10]), which confirms Medline as the most efficient bibliographic database for retrieving original articles in the biomedical informatics field.

The optimized query didn't retrieve articles yet included in the reviews, 1 in [2], 3 in [9], and 5 in [10]. Taking into account that the article in [2] was also included in [9], the optimized query missed 8 articles. For all these 8 articles, the term *nurse* or its variation was absent in the title and in the abstract (e.g., more general terms were utilized like healthcare workers). The article found in [2] and [9] should have been retrieved in the RL, and will be added to the final list of results. When refining the query (to get from 1581 returned articles to 3283), we tried to include the relevant missing articles. However, we stuck to a query we wanted to be focused on CDSSs for nursing practices taking the risk to miss some articles (mainly because of indexing errors in such articles). The remaining seven articles were out of the scope (and will not be added to RL).

Table 1. Precision and recall for the 3 selected reviews computed from papers published in [2014;2017], articles retrieved from BDBs and Medline (M), included in RL and included in RL but not retrieved by the review.

Review	#returns in BDBs/M	# articles in BDBs/M	# articles [2014-2017] in BDBs/M	# articles [2014-2017] in RL and in review	# articles [2014-2017] in RL and not in review	Precision on M	Recall on M
Akbar *et al.* [2]	1019/141	28/25	17/15(12)	11	27	91.70%	28.90%
Borum [9]	29/ NA	9/9	8/8	5	2	62.50%	71.40%
Mebrahtu *et al.* [10]	49,852/ 9,549	35/33	9/8	3	8	37.50%	27.30%

When performing a specific Medline search, we cannot foresee to what extent the search will be successful. Calculating precision and recall can help in understanding and tuning the search strategy. Therefore, when starting a systematic review, it is of major importance to have some sense of how the query is performing. Given the inverse relationship between precision and recall, unusually high precision rate may be indicative of overly restrictive searching at the expense of the complete identification of relevant studies (recall). For instance, the review of Akbar *et al.*, [2] was interested in studying the effects of CDSSs solely used by nurses on decision-making, care delivery, and patient outcomes. Despite the large scope and the number of outcomes included, it achieved a high precision and a low recall. This can be related to missing synonyms and controlled vocabulary (MeSH terms) in the query, narrowing the results to only 141 articles from Medline despite many outcomes were searched.

Having almost the same scope, Mebrahtu *et al.* [10] studied the effects of CDSSs on nursing and *allied* health professional performance, and on patient outcomes. This review has low precision and low recall. However, the initial query for this review was made of very broad terms and the authors deliberately aimed at a high recall (49,852 returns). Reviewers must be prepared to maintain rigorous screening standards in the face of large search retrievals. It is also important to know that Mehrabtu *et al.* [10] limited the types of studies to include only methodologically sound studies (such as randomized controlled trials), which may explain why we only added 8 studies from the RL, as compared to Akbar *et al.* [2] where we added 27 articles.

Borum [9] has a very specific scope focused on the barriers to the use of CDSSs perceived by nurse practitioners in hospital settings. The high recall may explain why we didn't find a significant number of missing articles regarding this topic.

Some limitations of our work are that we only assessed Medline database, we restrained the time frame to [2014;2017], we only considered three reviews, and we restricted the search to the literature on CDSSs for nursing practice. In addition, we proposed a proxy measure for precision and recall assuming the RL was the gold standard. Thus, true positives in our work may be different from those considered as true positives in the three reviews. Lastly, it is important to take into consideration that missing references in reviews may be due to factors beyond the control of reviews' authors. Indexing errors (coming from missing MeSH terms, or forgotten common terms in title or abstract) may lead to make relevant articles unavailable to systematic reviews.

5. Conclusion

Errors in queries are present in the majority of reviews which may limit recall and precision of search strategies. A good compromise between the two measures should be sought, in order to reach an acceptable tradeoff between result comprehensiveness and resource utilization/workload. This is why, before conducting a literature review in the domain of computerized decision support systems in nursing, we wanted to develop a comprehensive search strategy by identifying the necessary concepts, finding MeSH terms and synonyms, avoiding syntax errors, and comparing and gauging the query to the body of existing knowledge and fine tuning it in order to have a comprehensive return.

While waiting for machine learning techniques to automate the screening of relevant studies while conducting systematic reviews, human effort would be required for the foreseeable future.

References

[1] Piscotty R, Kalisch B. Nurses' use of clinical decision support: A literature review. Vol. 32, CIN - Computers Informatics Nursing. 2014. doi: 10.1097/CIN.0000000000000110.

[2] Akbar S, Lyell D, Magrabi F. Automation in nursing decision support systems: A systematic review of effects on decision making, care delivery, and patient outcomes. Vol. 28, Journal of the American Medical Informatics Association. 2021. doi: 10.1093/jamia/ocab123

[3] Salvador-Oliván JA, Marco-Cuenca G, Arquero-Avilés R. Errors in search strategies used in systematic reviews and their effects on information retrieval. Journal of the Medical Library Association. 2019;107(2). doi: 10.5195/jmla.2019.567

[4] Straube S, Heinz J, Landsvogt P, Friede T. Recall, precision, and coverage of literature searches in systematic reviews in occupational medicine: an overview of Cochrane Reviews. GMS Medizinische Informatik, Biometrie und Epidemiologie [Internet]. 2021 [cited 2023 Jan 8];17(A). Available from: https://www.egms.de/static/en/journals/mibe/2021-17/mibe000216.shtml

[5] Anderson JA, Willson P. Clinical decision support systems in nursing: Synthesis of the science for evidence-based practice. CIN - Computers Informatics Nursing.2008;26(3).doi: 10.1097/01.NCN.0000304783.72811.8e.

[6] Randell R, Mitchell N, Dowding D, Cullum N, Thompson C. Effects of computerized decision support systems on nursing performance and patient outcomes: A systematic review. Vol. 12, Journal of Health Services Research and Policy. 2007. doi: 10.1258/135581907782101543.

[7] Lee S. Features of computerized clinical decision support systems supportive of nursing practice: A literature review. Vol. 31, CIN - Computers Informatics Nursing. 2013. doi: 10.1097/01.NCN.0000432127.99644.25.

[8] Dunn Lopez K, Gephart SM, Raszewski R, Sousa V, Shehorn LE, Abraham J. Integrative review of clinical decision support for registered nurses in acute care settings. Vol. 24, Journal of the American Medical Informatics Association : JAMIA. 2017. doi: 10.1093/jamia/ocw084.

[9] Borum C. Barriers for Hospital-Based Nurse Practitioners Utilizing Clinical Decision Support Systems: A Systematic Review.Vol. 36,CIN - Computers Informatics Nursing.2018.doi: 10.1097/CIN.0000000000000413.

[10] Mebrahtu TF, Skyrme S, Randell R, Keenan AM, Bloor K, Yang H, et al. Effects of computerised clinical decision support systems (CDSS) on nursing and allied health professional performance and patient outcomes: A systematic review of experimental and observational studies. Vol. 11, BMJ Open. 2021. doi: 10.1136/bmjopen-2021-053886.