Published in final edited form as:

J Clin Epidemiol. 2015 April ; 68(4): 412–417. doi:10.1016/j.jclinepi.2014.10.008.

Citation searches are more sensitive than keyword searches to identify studies using specific measurement instruments

Suzanne K. Linder^{1,2}, Geetanjali R. Kamath¹, Gregory F. Pratt³, Smita S. Saraykar¹, and Robert J. Volk¹

¹Department of General Internal Medicine, The University of Texas MD Anderson Cancer Center, Houston, TX

²Department of Rehabilitation Sciences, Sealy Center on Aging, The University of Texas Medical Branch, Galveston, TX

³Research Medical Library, The University of Texas MD Anderson Cancer Center, Houston, TX

Abstract

Objective—To compare the effectiveness of two search methods in identifying studies that used the Control Preferences Scale (CPS), a healthcare decision-making instrument commonly used in clinical settings.

Study Design & Setting—We searched the literature using two methods: 1) keyword searching using variations of “control preferences scale” and 2) cited reference searching using two seminal CPS publications. We searched three bibliographic databases [PubMed, Scopus, Web of Science (WOS)] and one full-text database (Google Scholar). We report precision and sensitivity as measures of effectiveness.

Results—Keyword searches in bibliographic databases yielded high average precision (90%), but low average sensitivity (16%). PubMed was the most precise, followed closely by Scopus and WOS. The Google Scholar keyword search had low precision (54%) but provided the highest sensitivity (70%). Cited reference searches in all databases yielded moderate sensitivity (45–54%), but precision ranged from 35–75% with Scopus being the most precise.

Conclusion—Cited reference searches were more sensitive than keyword searches, making it a more comprehensive strategy to identify all studies that use a particular instrument. Keyword

Corresponding Author: Suzanne K. Linder, PhD, 301 University Blvd, Galveston, TX 77555-0177, T 409-266-9647, sulinder@utmb.edu.

Author contributions

SL and RV conceptualized the paper. SL, GK, GP, SS and RV planned the study design. SL, GK, GP and SS collected and analyzed the data presented in the paper. SL and SS wrote the initial draft of the paper. SL, GK, GP, SS and RV interpreted the study results, added to and critically revised the paper for intellectual content. All authors approved the final version of the paper.

Competing interests: All authors declare: no support from any organization for the submitted work; no financial relationships with any organizations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

searches provide a quick way of finding some but not all relevant articles. Goals, time and resources should dictate the combination of which methods and databases are used.

Keywords

search methods; information sources; meta-analyses; systematic reviews; instruments; methodology

INTRODUCTION

Systematic reviews with meta-analyses assembled from well-designed and well-executed primary studies are regarded as the highest level of evidence. Good quality reviews are of great importance to clinicians practicing evidence-based medicine [1] (i.e., use of the available best evidence to make decisions about treatment of individual patients or delivery of healthcare services); and evidence-based clinical practice [2] (i.e., presenting the best evidence to patients and consulting with them in order to select the most suitable treatment option).

A thorough search of the literature is a key component of these reviews, and may be conducted by authors or librarians. Most commonly, the searcher constructs a topical search strategy using a combination of keywords and database-specific subject heading terms and combines them in ways that maximize the likelihood of identifying all relevant studies.

However, when the purpose of a systematic review is to compare an outcome across studies that use the same assessment instrument, the task can be difficult to achieve with a topical literature search. Instruments are not well-indexed as subject headings in the major biomedical databases and the instrument name is usually not present elsewhere in the database record (e.g. title or abstract). No link exists between the database record and the instrument used in the study.

Furthermore, searching with the outcome name as part of the topical search may fail to locate studies using the same instrument. Authors may use different terms for the outcome across studies. Additionally, the outcome may not be mentioned in the abstract if it was not the primary one for the study.

Recently, systematic review methodology research has focused on determining effective search strategies beyond the traditional methods. New technologies for cross-database searching and the development of full-text databases have increased the complexity of searching and have added new types of search methods [3]. Current research indicates that reliance on one search method or process may yield different results than another method or process, affecting the number of studies found and analyzed [4, 5]. This includes methods or strategies that use different types of search terms [4, 6–8], different search techniques [5] including the use of search filters [9, 10], and the number of databases searched [11–17]. Therefore, using a combination of search strategies, techniques, and multiple databases is recommended for most systematic reviews in order to identify all relevant articles.

One search technique that deserves more attention is cited reference searching. A cited reference search begins with identifying an original article and then searching for newer articles that cite that original paper. Papaioannou and colleagues [5] found citation searches to be an effective supplement to traditional subject searches to locate additional high-quality references for a systematic review on student experience of e-learning in the social sciences literature. However, Wright and colleagues [18] found citation searching to not add any additional references for a systematic review of interventions for multiple risk behaviors after completing a traditional subject search. The discrepancy in the effectiveness of citation searching between these two studies is probably due to the topic and goal of the systematic reviews.

Cited reference searches have the potential to be an effective method in retrieving studies that use the same assessment instrument. It may be more effective than keyword searches due to the common practice among authors to acknowledge the instrument developers by citing the first publication of the instrument, the validation study, or both articles when describing measures used in studies.

Our overall goal of this study was to compare the effectiveness of keyword and cited reference search methodologies in the retrieval of references for a systematic review that compared patient participation in decision making across a variety of healthcare decisions (e.g. cancer screening and treatment, chronic diseases). For this case study, we selected the Control Preferences Scale (CPS) as the target of our keyword and cited reference literature searches. The scale is a validated, self-report instrument commonly used to assess desired involvement or participation in healthcare decision making [19, 20]. It has been included as an outcome measure in several systematic reviews [21–26] including a review that focused solely on findings from the CPS [27]. We hypothesized that cited reference searches would be more sensitive than keyword searches to identify studies that report CPS findings.

A case study to compare search methods

Information sources and search process—We searched for studies using the CPS in four databases: Google Scholar, Scopus, and Web of Science, and PubMed. Three of these (Google Scholar, Web of Science, and Scopus) have the ability to perform both types of searches. Google Scholar is a full text database and the other two are bibliographic (i.e. database records are created as searchable surrogates of complete articles). We also included PubMed, because it is the most commonly searched bibliographic database for biomedical topics and is freely available.

To standardize our search results, we limited retrieval for both types of searches to the ten-year time period of 2003–2012. For the keyword searches, we searched for “control preference scale” OR “control preferences scale” (i.e. as exact word phrases) in all four databases. In addition to being a common practice for topical searches, using the exact word phrase standardized the search in all four databases and eliminated searcher expertise and search system capabilities as confounding factors when comparing search methods.

For the cited reference searches, we used two seminal publications: 1) the initial article that introduced the CPS published in 1992 [19], and 2) the validation study that was published in

1997 [20]. We believed a high percentage of studies using the CPS would cite one or both of these. In Google Scholar, we performed separate cited reference searches for the 1992 and 1997 articles. This enabled comparing Google Scholar results to Web of Science for the 1992 article (Web of Science does not cover the journal that published the 1997 article) and to Scopus for the 1997 article (Scopus records begin about 1996 so there is no record for the 1992 article). Table 1 provides additional details about the databases and how each was searched.

Comparison of search methods—We defined “relevant” articles as those in which the CPS was used as an assessment instrument and determined this by examining the full text. Articles that cited one of the two references but did not use the CPS to gather data were “non-relevant.” To compare the search methods (keyword and cited reference searching) for information retrieval effectiveness, we assessed precision and sensitivity (recall) [28, 29]. We calculated *precision* for a database search as the percentage of the relevant articles relative to the total number of citations retrieved in that search. We calculated *sensitivity* for a database search as the percentage of relevant articles relative to the total number of unique relevant articles when all search results were combined (i.e. keyword and cited reference searching of all databases combined, accounting for duplicate citations).

Findings

After discarding non-English language and duplicate citations, the combined keyword and cited reference searches retrieved 740 unique citations. The CPS was used in 277 of these (37%). Figure 1 presents precision and sensitivity data for each search method by database.

Keyword searching of the three bibliographic databases (PubMed, Scopus, and Web of Science) provided high precision, averaging 90%, but an average sensitivity of only 16%. If the CPS was mentioned in the database record, there was a very good chance it was used in the study, but many CPS studies were absent from each set of results. Keyword searching of Google Scholar provided contrasting results in that sensitivity was high at 70% (193 of 277 possible CPS studies) but precision was markedly lower than with the bibliographic databases at 53%. We assessed 365 Google Scholar items to find 193 CPS studies.

For the four cited reference searches, sensitivity was similar ranging from 45–54%. Approximately one-half of the 277 CPS studies were found regardless of which database or which reference served as the starting point. The precision of these four cited reference searches was more varied ranging from 35% to 75%. Searches using the 1997 validation study as the starting point (Google Scholar and Scopus) averaged nearly 70% precision. When the 1992 seminal CPS article was the starting point (in Google Scholar and Web of Science), precision averaged slightly less than 40%.

In both Scopus and Web of Science, cited reference searching found approximately three times as many studies using the CPS as were found by keyword searching in these databases. The tradeoff was that with a cited reference search, about two records needed to be assessed to find one using the CPS. With the keyword searches, this correspondence was much closer to one-to-one.

The two cited reference searches in Google Scholar (with the 1992 and 1997 articles as starting points) provided interesting results. Each search identified about 150 studies where the CPS was used, but the precision was 35% for the 1992 article search and 63% for the 1997 article search. The 1992 article was referenced 429 times and the 1997 article 237. These results were similar to those of Scopus and Web of Science and could indicate that a validation study is the more precise reference to begin with or merely that older studies have more time to be discussed in articles that are not research studies.

Our results indicate that regardless of the searching method chosen, it is critical to search more than one database when it is important to find a large proportion of available studies. Only 18 of the 740 unique citations were found in every search, a little over 2%. Of the 277 CPS studies, approximately 14% were found in all of the cited reference searches or all of the keyword searches. Only 17 CPS studies (6%) were found in the results of every search.

Based on our analysis of these results, we suggest a cited reference search in the Scopus database may be the most efficient place to start when trying to identify studies that used a specific instrument. Scopus had the highest cited reference search precision at 75% and sensitivity of 52% nearly matched that of Google Scholar (54%).

Both Scopus and Web of Science require subscriptions, and user access is often through academic or medical library affiliation. If only free databases (PubMed and Google Scholar) are available, a keyword search of Google Scholar may be the place to begin when trying to identify a large number of studies using a particular instrument. In our case study, this search provided 70% sensitivity (193 of 277 CPS studies) with 53% precision compared to the cited reference searches of Google Scholar that averaged 54% sensitivity and 49% precision. PubMed provided high precision (93%) for the keyword search, but found only 43 of the 277 CPS studies for a sensitivity of 16%.

We believe this is the first study to compare cited reference and keyword searching to identify studies using a particular instrument. In 2008, Falagas and colleagues [15] compared the same four databases used in our study for their content and utility characteristics and presented the advantages and disadvantages of each. They used both a keyword search for “brucellosis” and searched for articles citing a particular recent article (i.e., within three years) related to brucellosis. Although they did not report precision and sensitivity, their cited reference search results were similar to ours in that Scopus provided more retrieval than Web of Science, and Google Scholar provided proportionately more “false positive” records than either of these two bibliographic databases.

A major strength of our study is that we evaluated more than 1500 citations to identify the 740 unique, English-language ones and then examined the methods sections and other portions of articles to determine which of these actually used the CPS. This provided a large sample size and a means for calculating sensitivity. When assessing search strategies and search results, sensitivity is difficult to measure, because it requires knowing the number of relevant records within the larger pool of records searched.

Our study also provides value by using identical means to compare results obtained from searching a full text database (Google Scholar) to results from searching bibliographic

databases. Recent literature debates the use of Google Scholar for systematic reviews [12–16]. Reported criticisms of Google Scholar include high numbers of results accompanied by low precision. Our results support this conclusion for keyword searches of Google Scholar and suggest cited reference searching may be imprecise as well. The degree of precision in cited reference searches ranged from 35%–63% and the level of sensitivity was comparable to our results for Scopus and Web of Science.

Our case study has limitations. The search results for the CPS may not be applicable to similar searches for other instruments. Other instruments may be more or less likely to be referred to in database records or article references. Results may also vary for a different information retrieval goal such as searching for a specific outcome reported in an instrument. In addition, our keyword search strategy was limited to the name of the instrument (and a variation of the name – “preferences” instead of “preference”) so that an identical search could be performed in each database. A more complex strategy that uses terms related to the instrument (i.e., role preferences, patient involvement) could provide different results.

Authors and database developers can both help to improve topical keyword searches for systematic reviews. We recommend that authors should mention the scale or instrument used in their study in the abstract in order to make the phrase more searchable. Similarly, database developers should make assessment instruments or scales a part of indexing methodology. Parallel recommendations have been made by others to improve searches that have different goals than our case study. For example, Gorecki and colleagues [6] believe that better indexing of subject-specific outcomes and the methodology used would be beneficial to help identify articles. Shaw and colleagues [8] have discussed the need for better methods to index qualitative research in order to identify articles that use qualitative and mixed methods designs.

Our findings may be of interest to methodologists who conduct systematic reviews to locate the measurement properties of instruments such as reliability and internal consistency. Terwee and colleagues [30] developed a PubMed search filter to more efficiently locate studies that report measurement properties of scales and instruments. A cited reference search may also be helpful to identify studies that report the psychometric properties of the measure used, especially when the measurement findings are not the primary focus of the paper. Future research is needed to compare cited reference searches to other search techniques in order to determine the most effective and efficient combination of strategies.

In conclusion, our findings support our hypothesis in part. When trying to identify many articles using a particular instrument, cited reference searching in Scopus or Web of Science is preferable to keyword searching in these resources or in PubMed/Medline. However, in Google Scholar, the only full-text database we searched, a keyword search for CPS found more studies than the cited reference searches and precision was slightly higher. If the goal is to identify a few studies that use a specific instrument, keyword searches of a bibliographic database may be the most efficient. When it is important to conduct a comprehensive search to maximize the number of studies found, both cited reference and keyword searches should be conducted in more than one database. We believe these findings

may aid the decision making of those who perform systematic reviews, particularly when outcomes of interest are often reported using specific instruments.

Acknowledgments

Funding: This project was funded in part by a grant from the National Institute of Health/National Cancer Institute (#R21 CA132669) and by a grant from the Agency for Healthcare Research and Quality (#R24 HS022134). The content is solely the responsibility of the authors and does not represent the official views of the National Institute of Health or the Agency for Healthcare Research and Quality.

References

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. Clinical orthopaedics and related research. 2007; 455:3–5. [PubMed: 17340682]
2. Gray, JAM. Evidence-based healthcare: how to make health policy and management decisions. London: Churchill Livingstone; 1997.
3. Lefebvre C, Glanville J, Wieland LS, Coles B, Weightman AL. Methodological developments in searching for studies for systematic reviews: past, present and future? Systematic reviews. 2013; 2:78. [PubMed: 24066664]
4. Egan M, Maclean A, Sweeting H, Hunt K. Comparing the effectiveness of using generic and specific search terms in electronic databases to identify health outcomes for a systematic review: a prospective comparative study of literature search methods. BMJ open. 2012;2.
5. Papaioannou D, Sutton A, Carroll C, Booth A, Wong R. Literature searching for social science systematic reviews: consideration of a range of search techniques. Health information and libraries journal. 2010; 27:114–22. [PubMed: 20565552]
6. Gorecki CA, Brown JM, Briggs M, Nixon J. Evaluation of five search strategies in retrieving qualitative patient-reported electronic data on the impact of pressure ulcers on quality of life. Journal of advanced nursing. 2010; 66:645–52. [PubMed: 20423399]
7. Jenuwine ES, Floyd JA. Comparison of Medical Subject Headings and text-word searches in MEDLINE to retrieve studies on sleep in healthy individuals. Journal of the Medical Library Association: JMLA. 2004; 92:349–53. [PubMed: 15243641]
8. Shaw RL, Booth A, Sutton AJ, Miller T, Smith JA, Young B, et al. Finding qualitative research: an evaluation of search strategies. BMC medical research methodology. 2004; 4:5. [PubMed: 15070427]
9. Cooper C, Levay P, Lorenc T, Craig GM. A population search filter for hard-to-reach populations increased search efficiency for a systematic review. Journal of clinical epidemiology. 2014; 67:554–9. [PubMed: 24698297]
10. McKibbon KA, Lokker C, Wilczynski NL, Haynes RB, Ciliska D, Dobbins M, et al. Search filters can find some but not all knowledge translation articles in MEDLINE: an analytic survey. Journal of clinical epidemiology. 2012; 65:651–9. [PubMed: 22424986]
11. Lemeshow AR, Blum RE, Berlin JA, Stoto MA, Colditz GA. Searching one or two databases was insufficient for meta-analysis of observational studies. Journal of clinical epidemiology. 2005; 58:867–73. [PubMed: 16085190]
12. Gehanno JF, Rollin L, Darmoni S. Is the coverage of Google Scholar enough to be used alone for systematic reviews. BMC medical informatics and decision making. 2013; 13:7. [PubMed: 23302542]
13. Giustini D, Boulos MN. Google Scholar is not enough to be used alone for systematic reviews. Online journal of public health informatics. 2013; 5:214. [PubMed: 23923099]
14. Boeker M, Vach W, Motschall E. Google Scholar as replacement for systematic literature searches: good relative recall and precision are not enough. BMC medical research methodology. 2013; 13:131. [PubMed: 24160679]
15. Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. FASEB journal: official publication of

the Federation of American Societies for Experimental Biology. 2008; 22:338–42. [PubMed: 17884971]

16. Bramer WM, Giustini D, Kramer BM, Anderson P. The comparative recall of Google Scholar versus PubMed in identical searches for biomedical systematic reviews: a review of searches used in systematic reviews. *Systematic reviews*. 2013; 2:115. [PubMed: 24360284]
17. Whiting P, Westwood M, Burke M, Sterne J, Glanville J. Systematic reviews of test accuracy should search a range of databases to identify primary studies. *Journal of clinical epidemiology*. 2008; 61:357–64. [PubMed: 18313560]
18. Wright K, Golder S, Rodriguez-Lopez R. Citation searching: a systematic review case study of multiple risk behaviour interventions. *BMC medical research methodology*. 2014; 14:73. [PubMed: 24893958]
19. Degner LF, Sloan JA. Decision making during serious illness: what role do patients really want to play? *Journal of clinical epidemiology*. 1992; 45:941–50. [PubMed: 1432023]
20. Degner LF, Sloan JA, Venkatesh P. The Control Preferences Scale. *The Canadian journal of nursing research = Revue canadienne de recherche en sciences infirmieres*. 1997; 29:21–43. [PubMed: 9505581]
21. Stacey D, Legare F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. *The Cochrane database of systematic reviews*. 2014; 1:CD001431. [PubMed: 24470076]
22. Tariman JD, Berry DL, Cochrane B, Doorenbos A, Schepp K. Preferred and actual participation roles during health care decision making in persons with cancer: a systematic review. *Annals of oncology: official journal of the European Society for Medical Oncology/ESMO*. 2010; 21:1145–51. [PubMed: 19940010]
23. Chewning B, Bylund CL, Shah B, Arora NK, Gueguen JA, Makoul G. Patient preferences for shared decisions: a systematic review. *Patient education and counseling*. 2012; 86:9–18. [PubMed: 21474265]
24. Hubbard G, Kidd L, Donaghy E. Preferences for involvement in treatment decision making of patients with cancer: a review of the literature. *European journal of oncology nursing: the official journal of European Oncology Nursing Society*. 2008; 12:299–318. [PubMed: 18486552]
25. Gaston CM, Mitchell G. Information giving and decision-making in patients with advanced cancer: a systematic review. *Social science & medicine*. 2005; 61:2252–64. [PubMed: 15922501]
26. Kiesler DJ, Auerbach SM. Optimal matches of patient preferences for information, decision-making and interpersonal behavior: evidence, models and interventions. *Patient education and counseling*. 2006; 61:319–41. [PubMed: 16368220]
27. Singh JA, Sloan JA, Atherton PJ, Smith T, Hack TF, Huschka MM, et al. Preferred roles in treatment decision making among patients with cancer: a pooled analysis of studies using the Control Preferences Scale. *The American journal of managed care*. 2010; 16:688–96. [PubMed: 20873956]
28. Reitz, JM. Online dictionary for library and information science. *Libraries Unlimited*; 2014.
29. Higgins, J.; Green, S., editors. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. The Cochrane Collaboration; 2011.
30. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*. 2009; 18:1115–23.

What is new?

- Several studies have compared the effectiveness of different databases to retrieve relevant studies – we expand on this by starting with the search method and then quantitatively comparing individual database results.

Key findings

- When searching in Scopus or Web of Science, cited reference searches were preferable to keyword searches for identifying all studies that have used a particular instrument.
- When searching Google Scholar, a full-text database, the keyword search found more studies and had slightly higher precision than cited reference search.

Implications

- If the goal is to identify a few studies that use a specific instrument, keyword searches of a bibliographic database may be sufficient.
- When it is important to provide a comprehensive list of studies using a specific instrument, both cited reference and keyword searches should be conducted in more than one database.

Figure 1.

Precision and sensitivity for keyword and cited reference search methods by database

Table 1

Description of search methods and search strategies by information source

Database	Database Type	Search Method(s) Available	Search Type(s) Used
Google Scholar	Full text	Keyword	“control preference scale” or “control preferences scale” as phrase in title or abstract
		Cited references	1992 reference cited
			1997 reference cited
PubMed	Bibliographic	Keyword	“control preference scale” or “control preferences scale” as phrase in title or abstract
Scopus	Bibliographic	Keyword	“control preference scale” or “control preferences scale” as phrase in title or abstract
		Cited reference	1997 reference cited
Web of Science	Bibliographic	Keyword	“control preference scale” or “control preferences scale” as phrase in title or abstract
		Cited reference	1992 reference cited