

WorldCat Local and the Clinician: Is Discovery Improved?

Linda C.K. Crook

Washington State University

University of Pittsburgh

Abstract

This study examines how the next-generation discovery system WorldCat Local (www.oclc.org/worldcatlocal/) impacts discovery of biomedical literature. The system emphasizes single-box keyword searching and the use of facets for more focused search queries. One purpose of this study is to determine if browsing results over advanced searching is a better discovery method for the busy clinician. This study also explores the effectiveness of keyword searching in WorldCat Local for retrieving a subset of widely cited clinical literature.

Introduction

One of the benefits of next-generation discovery systems is the ability to integrate article-level records for metasearching (OCLC, 2010b). Specifically, the inclusion of records from PubMed in WorldCat.org's central index allows biomedical article resources to be searched along with traditional library catalog records. Since many users seem to prefer single-box keyword searches (Judkins, 2010), appropriate for use in WorldCat Local, this study sought to investigate whether busy clinicians conducting keyword searches were likely to find the materials they needed.

Methods

Several studies comparing PubMed with other databases or comparing different interfaces for PubMed and Medline were examined for this study, in search of a reproducible method that could build on previous research. Anders and Evans (2010) compared PubMed and Google Scholar by using cited literature from a Cochrane Collaboration systematic review as a benchmark for the search results. Falagas, Ntziora, Makris, Malietzis, & Rafailidis (2009) gave investigators case reports to diagnose and compared results of their PubMed and Google searches. Gall and Brahmi (2004) compared Medline searches in EndNote to direct searching in PubMed, using four search strings. Walters (2009) compared twelve databases using a single search string, judging recall and precision based on a pre-identified set of 155 documents.

Vanhecke, Barnes, Zimmerman, & Shoichet (2007) compared the ability of PubMed and HighWire Press to locate specific articles using a list developed by Ioannidis (2005). Ioannidis created this list of the 49 most highly cited clinical studies published between 1990 and 2003 in order to identify those with results that were contradicted or found less effective in later studies. For each of these articles, Ioannidis indicated "type of intervention and disease." Vanhecke et al.

(2007) used these brief descriptions as keywords in their comparative searches of PubMed and HighWire Press. They then measured the frequency of successfully locating the desired article within the first 25 or 200 results. Their purpose of simulating “real life usage by a clinician for ‘quick look-up,’” is a good match for the present study (Vanhecke, et al., 2007, p.1257). Further, the searches are replicable and the desired results were specific.

The methods developed in Vanhecke et al. (2007) are therefore adapted for use in the present study. To simulate user searching and avoid search bias, searches were performed based on the keywords from Ioannidis (2005). For purposes of expediency, the Year facet was used to focus results on the desired article. To verify that the specified articles were present in the database, secondary searches were run using the first four article title keywords, again using the year of publication to limit results. Although these comparison searches were originally run only for verification, they became an important source of comparative keyword data.

Searches were conducted in WorldCat.org in March and April 2011 using the default search option. The searches are saved in WorldCat.org so they can be run again as needed (see Figure 1 or <http://www.worldcat.org/profiles/lindackcrook/savedsearches>).

Results

In 17 of 49 searches (see Table 1), subject keywords failed to retrieve the desired article. The brief article descriptions by Ioannidis included 4-10 keywords each, most of which did not appear in the article citation. In some cases only a spelling change was needed to make the keyword searches successful, while some cases required a reduction in keywords. Searching with alternate spellings (especially in the case of hyphenated words) and synonyms would improve many of the searches.

In contrast, searches using the first four article title keywords (see Table 1) were successful in all but one case. A spelling variation between Ioannidis's (2005) "Alpha-2b" and the database's "Alpha2b" caused that search to fail.

Limitations

For all searches, resources retrieved in the first ten results—the default number displayed—seemed relevant, but this determination should ideally be made by clinicians, with clear standards for relevancy. That approach was beyond the scope of this study. An examination of what practitioners consider a successful search would be particularly revealing. In clinical settings, practitioners often do not search for a specific article, but for articles on a specific subject, to provide information for a current case. Searches that "failed" based on specific article retrieval tests might be considered successful in a discovery context.

The articles identified by Ioannidis (2005) are solely from major journals, and mostly from 1993. A more varied sample might be more illuminating.

Conclusions

Keyword searching can be successful in WorldCat Local both for finding specific articles and for finding materials by topic. Clinicians would be best served by limiting their starting keywords and varying their search terms.

The importance of recall, relevance, and precision must be re-examined in the next-generation discovery environment. Today's busy researcher or clinician has "imprecise search skills" (Haynes, 2010, p.448) and is likely to search using keywords only. Thiele, Scalzo, Pairo, & Nemergut (2010) based their comparison study on the assumption that "the average physician is unwilling to spend more than a few minutes searching for the information required to answer biomedical questions" (p.459). Recall and precision may therefore be less important than the

relevance of the first pages of results, which may provide adequate resources for the immediate need. Karimi, Zobel, Pohl, & Scholer (2009), in examining Boolean and ranked retrieval methods, determined, “recall is not as important as having a large number of relevant documents...sufficient to reward a searcher’s effort” (p.92). Determination of relevance and sufficiency can only be made by clinicians on a case-by-case basis. User testing is needed to fully evaluate next-generation discovery.

References

- Anders, M. E., & Evans, D. P. (2010). Comparison of PubMed and Google Scholar literature searches. *Respiratory Care*, 55(5), 578-83.
- Arcolio, A. & Davidson, S. (April 5, 2010). WorldCat Local at the University of California: Usability Testing: Round Two, Fall 2009: Findings and recommendations from the UC-OCLC fall 2009 assessment project : Summary report.
http://libraries.universityofcalifornia.edu/about/oclc_docs/NextGenMelylFindingsApril2010.pdf
- Arcolio, A. & Poe, F. (September 3, 2008). WorldCat Local at the University of California: Usability Testing: Round One, Spring 2008 : Findings and recommendations from the UC-OCLC spring 2008 assessment project : Summary report.
http://libraries.universityofcalifornia.edu/about/oclc_docs/WCL_Summary_03Sept2008.pdf
- Breeding, M. (2010). *Next-gen library catalogs*. New York: Neal-Schuman Publishers.
- Connor, E. (2010). PubMed search interface alternatives: A descriptive comparison. *Journal of Electronic Resources in Medical Libraries*, 7(2), 126-134.
- Dorman, D. (2008). The potential of metasearching as an "open" service. *Library Hi Tech*, 26(1), 58-67.
- Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *Faseb Journal*, 22 (2), 338-342.
- Falagas, M. E., Ntziora, F., Makris, G. C., Malietzis, G. A., & Rafailidis, P. I. (2009). Do PubMed and Google searches help medical students and young doctors reach the correct diagnosis? A pilot study. *European Journal of Internal Medicine*, 20(8), 788-790.
- Freeman, M. K., Lauderdale, S. A., Kendrach, M. G., & Woolley, T. W. (2009). Google scholar versus PubMed in locating primary literature to answer drug-related questions. *Annals of Pharmacotherapy*, 43(3), 478-484.
- Gall, C., & Brahmi, F. (2004). Retrieval Comparison of EndNote to Search MEDLINE (Ovid and Pub Med) versus Searching Them Directly. *Medical Reference Services Quarterly*, 23(3), 25-32.
- Haynes, R. B., Wilczynski, N., McKibbin, K. A., Walker, C. J., & Sinclair, J. C. (1994). Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *Journal of the American Medical Informatics Association*, 1(6).

- Ioannidis, J. P. (2005). Contradicted and initially stronger effects in highly cited clinical research. *The Journal of the American Medical Association*, 294(2), 218-28.
- Judkins, D. (2010, February). Expert searching: So you want to use Google. *MLA News*, 50(2), 14.
- Judkins, D. (2011). *So you want to use Google: Comparing PubMed, CINAHL, Google, and Google Scholar*. Poster session presented at the 2011 Medical Library Association Annual Conference, Minneapolis, MN.
- Karimi, S., Zobel, J., Pohl, S., & Scholer, F. (2009). The challenge of high recall in biomedical systematic search. *International Conference on Information and Knowledge Management, Proceedings*, 89-92.
- Kimball, R. (2010). The GeoRef database: A detailed comparison and analysis of three platforms. *Science and Technology Libraries*, 29, 111-129.
- Kimball, R. (2010). The TULSA petroleum abstracts database: A comparison of two platforms. *Science and Technology Libraries*, 29, 92-110.
- Mastrangelo, G., Fadda, E., Rossi, C.R., Zamprogno, E., Buja, A., & Cegolon, L. (2010). Literature search on risk factors for sarcoma: PubMed and Google Scholar may be complementary sources. *BMC Research Notes*, 3, 131.
- Meert, D., & Fitzgibbons, M. (2010). Are bibliographic management software search interfaces reliable?: A comparison between search results obtained using database interfaces and the EndNote online search function. *The Journal of Academic Librarianship*, 36(2), 144-150.
- Montori, V. M., Wilczynski, N. L., Morgan, D., Haynes, R. B., & Hedges Team. (2005). Optimal search strategies for retrieving systematic reviews from Medline: analytical survey. *BMJ (Clinical Research Ed.)*, 330, 7482.
- OCLC. (2010a). *Findings from WorldCat Local Usability Tests, July, 2009-September 2010*. http://www.oclc.org/worldcatlocal/about/213941usf_some_findings_about_worldcat_local_2011.pdf
- OCLC. (2010b). Metasearch expands the reach of WorldCat Local. *Next Space*, 14, 18-18.
- Shadle, S. (2009). Electronic resources in a next-generation catalog: The case of WorldCat Local. *Journal of Electronic Resources Librarianship*, 21, 192-199.
- Shultz, M. (2007). Comparing test searches in PubMed and Google Scholar. *Journal of the Medical Library Association*, 95(4), 442-5.

Spring, H. (2010). Health professionals of the future: Teaching information skills to the Google generation. *Health Information and Libraries Journal*, 27(2), 158-162.

Steinbrook, R. (2006). Searching for the right search--reaching the medical literature. *The New England Journal of Medicine*, 354(1), 4-7.

Thiele, R. H., Scalzo, D. C., Poiró, N. C., & Nemergut, E. C. (2010). Speed, accuracy, and confidence in Google, Ovid, PubMed, and UpToDate: Results of a randomised trial. *Postgraduate Medical Journal*, 86(1018), 459-465.

Vanhecke, T. E., Barnes, M. A., Zimmerman, J., & Shoichet, S. (2007). PubMed vs. HighWire Press: A head-to-head comparison of two medical literature search engines. *Computers in Biology and Medicine*, 37(9), 1252-8.

Walters, W. H. (2009). Google Scholar search performance: Comparative recall and precision. *Portal: Libraries and the Academy*, 9(1), 5-24.

WorldCat Local Known-Item Searching Focus Group, January 10, 2010: Notes. (2010).

Table 1

Keyword Searches

Search #	Keyword search	No. keywords	Item found	Item found w/title keywords
1	zidovudine asymptomatic HIV-1 infection	4	yes	yes
2	lipid lowering decrease coronary lesions coronary artery disease	8	no	yes
3	levamisole fluorouracil colon cancer	4	yes	yes
4	enalapril hydralazine isosorbide congestive heart failure	6	top hit	yes
5	postmenopausal hormonal therapy coronary artery disease prevention	7	no	yes
6	carotid endarterectomy high-grade stenosis	5	yes	yes
7	monoclonal antibody endotoxin gram-negative sepsis	6	yes	yes
8	enalapril patients left ventricular dysfunction	5	no	yes
9	captopril patients myocardial infarction	4	yes	yes
10	angioplasty tPA thrombolysis acute myocardial infarction	6	no	yes
11	captopril slowing disease progression diabetic nephropathy	6	no	yes
12	vitamin E coronary artery disease prevention men	7	no	yes
13	vitamin E coronary artery disease prevention women	7	no	yes
14	nitric oxide inhalation acute respiratory distress syndrome	7	no	yes
15	intensive management reduce type 1 diabetes complications	7	no	yes
16	7E3 high-risk angioplasty	4	no	yes
17	zidovudine reduce perinatal HIV-1 transmission	6	no	yes
18	stent balloon angioplasty coronary artery disease	6	top hit	yes
19	stent balloon angioplasty single-vessel coronary artery disease	8	no	yes
20	vitamin E beta carotene lung cancer	6	yes	yes
21	rt-PA acute stroke	4	yes	yes
22	pravastatin hypercholesterolemia	2	yes	yes
23	pravastatin myocardial infarction average cholesterol	5	top hit	yes
24	carvedilol congestive heart failure	4	no	yes
25	beta carotene retinol preventing lung cancer coronary artery disease	9	no	yes
26	aspirin prevent myocardial infarction men various C-reactive protein levels	10	no	yes
27	triple therapy indinavir 2 nucleosides hiv-1 infection	8	no	yes
28	abciximab glycoprotein IIb/IIIa blockade PCI	6	yes	yes
29	interferon alfa-2b ribavarin interferon alone chronic hepatitis C	9	no	yes
30	pravastatin secondary coronary artery disease prevention	6	no	yes
31	spironolactone severe congestive heart failure	5	no	yes
32	ramipril prevent coronary artery disease high-risk patients without left ventricular dysfunction congestive heart failure	15	no	yes
33	treatment systolic hypertension elderly adults	5	no	yes
34	postmenopausal estrogen progestin coronary artery disease risk factors	8	no	yes

Search #	Keyword search	No. keywords	Item found	Item found w/title keywords
35	endarterectomy asymptomatic stenosis >60%	4	yes	yes
36	estrogen progestin secondary coronary artery disease prevention	7	no	yes
37	lovastatin primary coronary artery disease prevention average cholesterol	8	no	yes
38	estrogen progestin coronary artery disease prevention	6	no	yes
39	folate prevent neural tube defects	5	no	yes
40	flavonoids coronary artery disease prevention	5	no	yes
41	simvastatin hypercholesterolemia previous coronary artery disease	6	no	yes
42	clopidogrel aspirin patients risk ischemic events	6	no	yes
43	vitamin E prevent myocardial infarction death patients coronary artery disease	10	no	yes
44	intensive blood-pressure lowering/low-dose aspirin hypertension	8	yes	yes
45	interferon alfa-2b ribavirin interferon alone chronic hepatitis C	9	yes	no
46	intensive management type 2 diabetes insulin sulphonylureas	7	no	yes
47	bisoprolol congestive heart failure	4	no	yes
48	all-trans retinoic acid acute promyelocytic leukemia	7	yes	yes
49	tamoxifen breast cancer prevention	4	yes	yes

Figure 1

Results of keyword searches as presented in WorldCat Local

